



INTERNATIONAL ASSOCIATION OF SCIENTIFIC, TECHNICAL & MEDICAL PUBLISHERS

---

# Looking at Metrics: Counting & What Counts

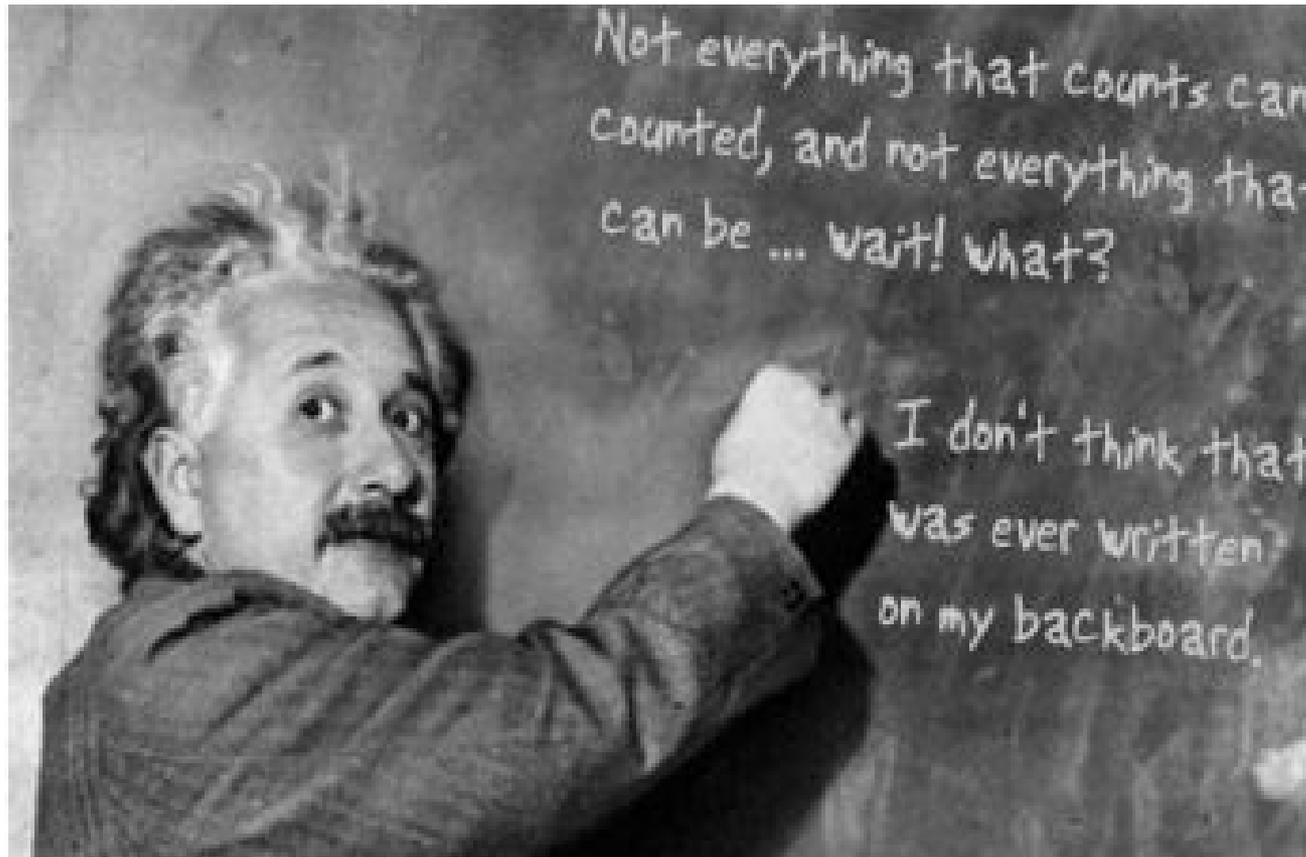
**Michael A Mabe**

*CEO, STM*

&

*Visiting Professor, Information Science,  
University College, London*

**Not everything that counts can be counted,.. not everything that can be counted counts — Einstein**



# On measurement

You can't manage  
what you can't **measure**

Peter Drucker



# Holy Grail of Research Assessment: Metrics!

- RA Objectives
  - Encourage excellence, ration funds
- RA Assumptions
  - $\alpha$  inputs directly lead to  $\alpha$  outputs
  - Institutions with most  $\alpha$  outputs are the best
  - Funding  $\alpha$  institutions leads to more  $\alpha$  outputs
  - More  $\alpha$  outputs are good value for society
- Metrics make this “easy”

# Research Assessment & Publishing

- Research assessment is at the heart of the scholarly system...
  - The awarding of **research grants**
  - The peer review of **papers**
  - The reputation of **journals**
- ... and publication activity helps judge...
  - The **standing of a researcher** from where s/he **publishes**; which feeds into
  - The **evaluation of institutions** based on the **publication** record of their staff
- Assessment is not an exact science

# How can it be done?

## “Beauty Contests & Footprints in the Sand”

- Qualitative expert opinion
  - Peer review: grant panels, journal referees etc.
- Quantitative metrics of outcomes
  - Mostly derivatives of publication
- Algorithm to combine and produce rankings
  
- Methodological issues
  - How to count metrics
  - How to normalise for comparison
  - Dealing with very different disciplinary profiles

# Pebbles in a pond



# Measurable Outcomes of Research

## OUTCOMES

- First order outcome
  - Publication
- Second order outcomes
  - Usage by readers
  - Citation by authors
  - Patents applied for
- Third order outcomes
  - Media coverage
  - Social media activity
- Fourth order outcomes
  - Development and economic exploitation

## METRICS

- Publication counts
- Download counts
- Citation counts
- Patents granted

**An issue for SSAH disciplines**

- No. of stories that appear
- Tweets etc.
- GDP and employment

# Normalisation Problems

- How/what/who do you count?
- Publication productivity
  - Per capita, per department, per discipline?
- Usage and citation
  - Downloads/cites per published item?
  - Downloads/cites per journal (or per e book)?
  - Downloads/cites per journal per market size?
  - Who (or what) is doing the downloading?
- Complex (compound) metrics
  - E.g., journal impact factors, how to associate with authors, institutions?

# Counting statistics

- Skewed distributions are the norm
- Data levels can be low at an article level
- Accuracy/error level depends on number of items ( $N$ ) in the count

$$\propto 1/\sqrt{N}$$

- so metrics with 100 data points will still report with an error of 10%, 400 5%, 900 3.3% etc.
- Distinguishing results from one publication unit to another may fall within statistical noise so all results are the same

# Citations As A Metric

- Assumption
  - Citation is an attribution of scholarly value
  - Not all fields are equally covered by citation datasets; does not work well where book publication predominates
- Wide range of citation types
  - Honorific, historic, “political”, discursive
  - Not all reflect scholarly value

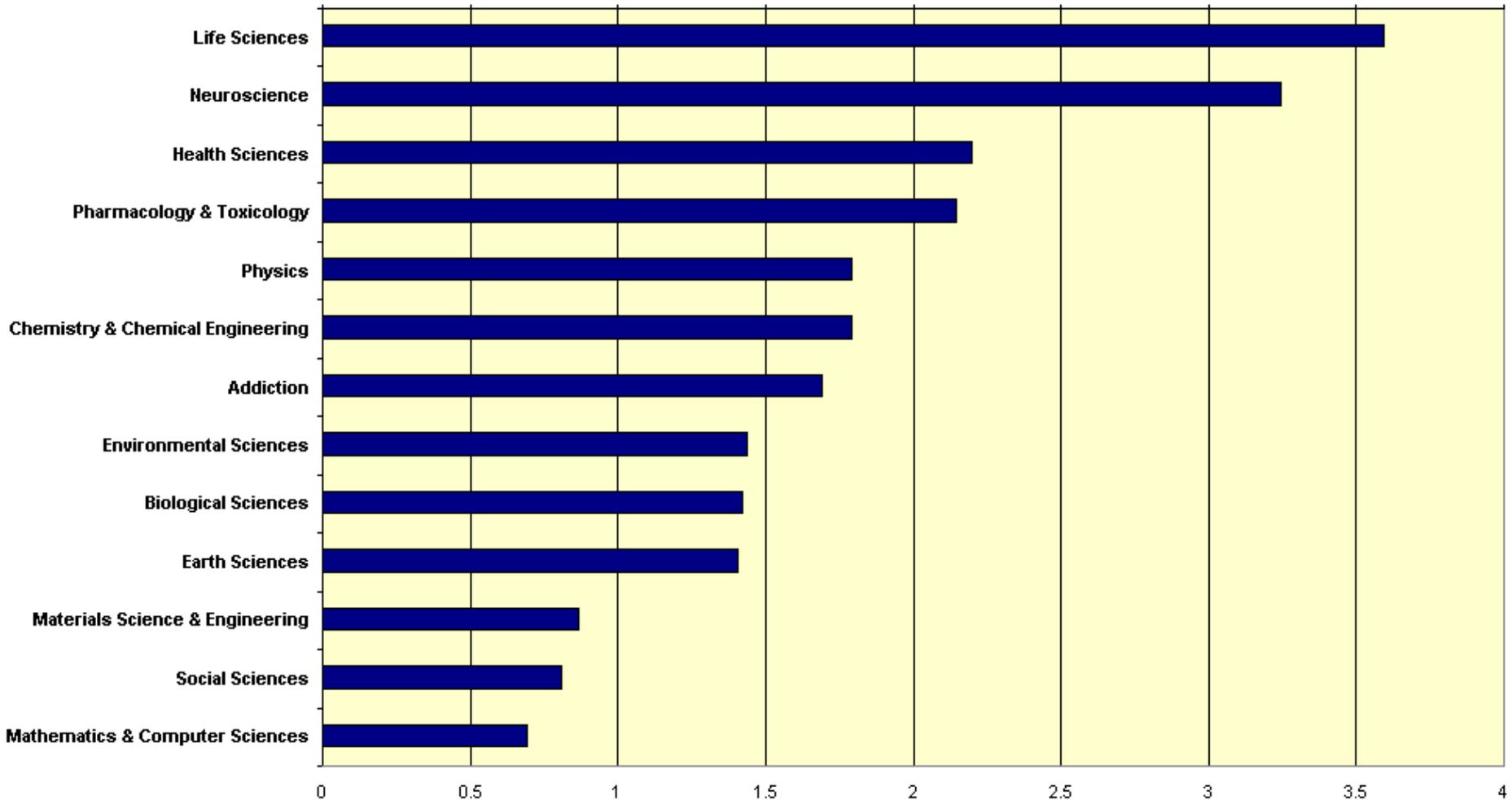
# Citations As A Metric

- Simple counting of citations is misleading
  - Journals have different sizes (papers/year), widely varying readerships, different purposes
  - Cross disciplinary comparison is invalid
- Analysis difficult at article or author level
  - Data is skewed: normal for 80% of cites to come from 20% of articles
  - Insufficient data: level close to statistical noise
  - All except minority of top papers (the 20% above) would be ranked the same

# Impact Factor

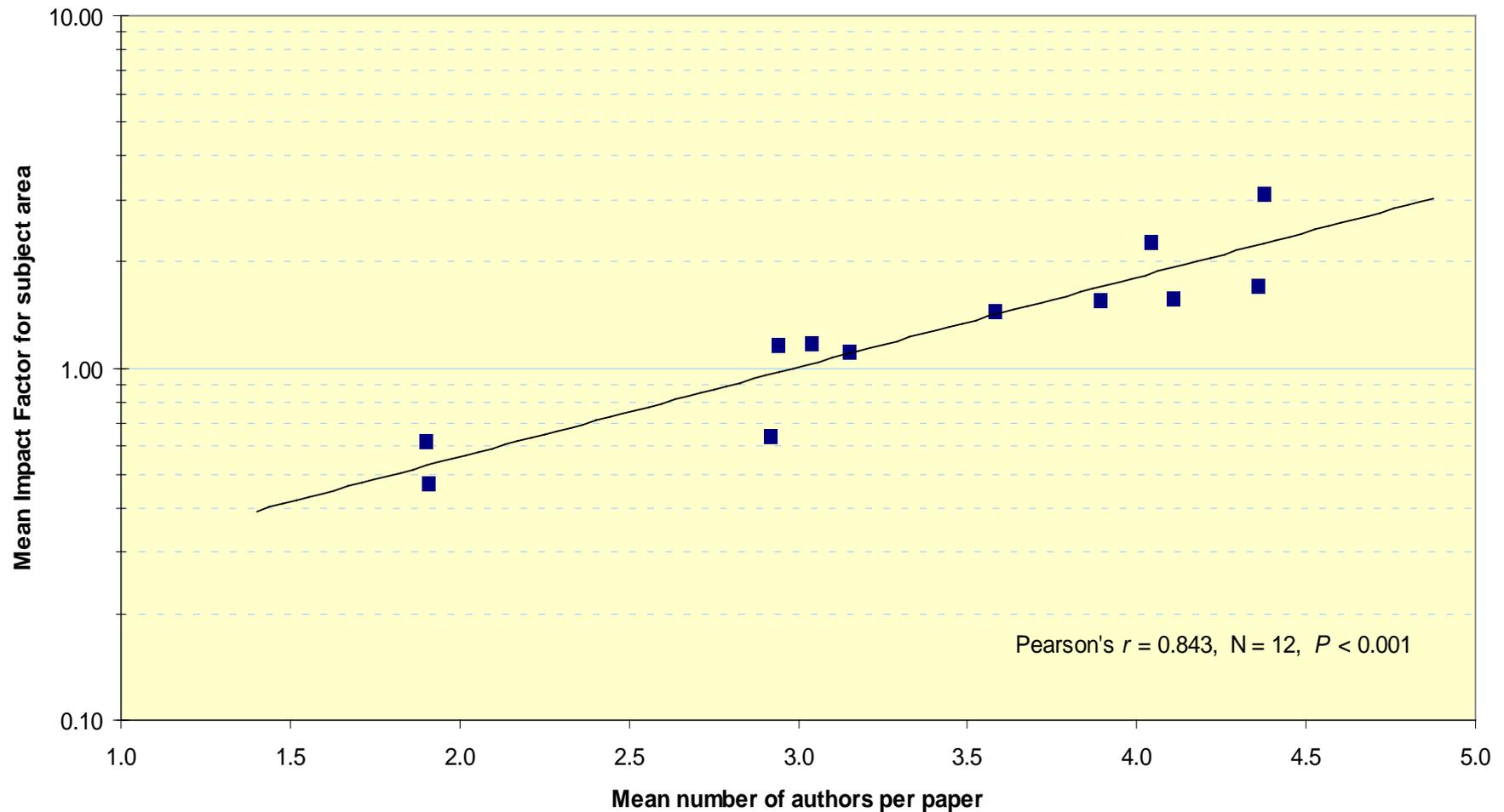
- A normalized citation measure
  - Number of citations received over a period  
divided by
  - number of citable items published in that period
- */S/* journal impact factor is over two years
  - Varies by subject area, journal (paper) type, journal size, citation window, and over time
  - Depends on definition of “citable items”
- **Cannot** be manipulated as a “score” or raw number

# Influences on Impact Factors: Subject Area

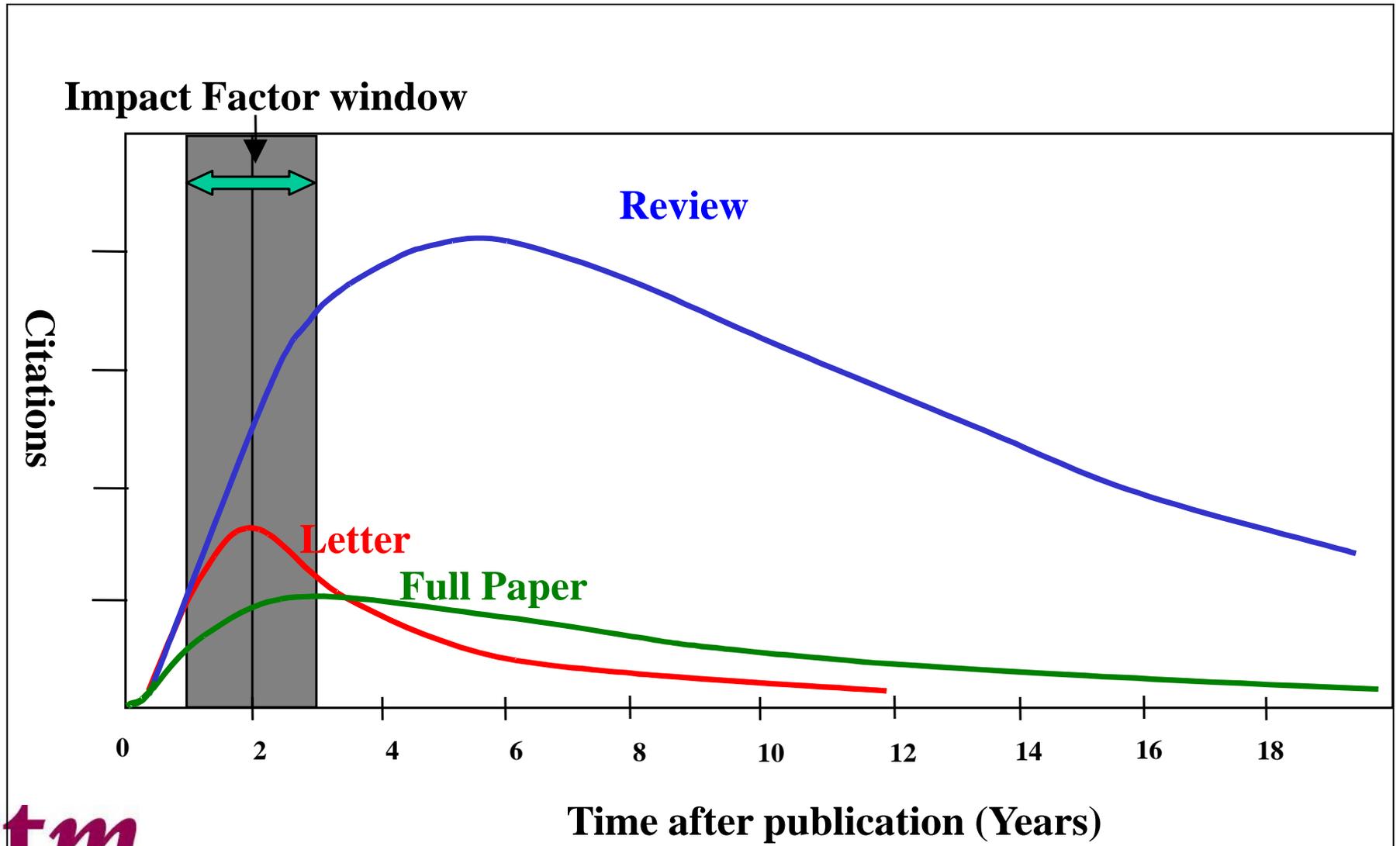


# Influences on Impact Factors: No of Authors

Impact Factors and Mean number of Authors per paper

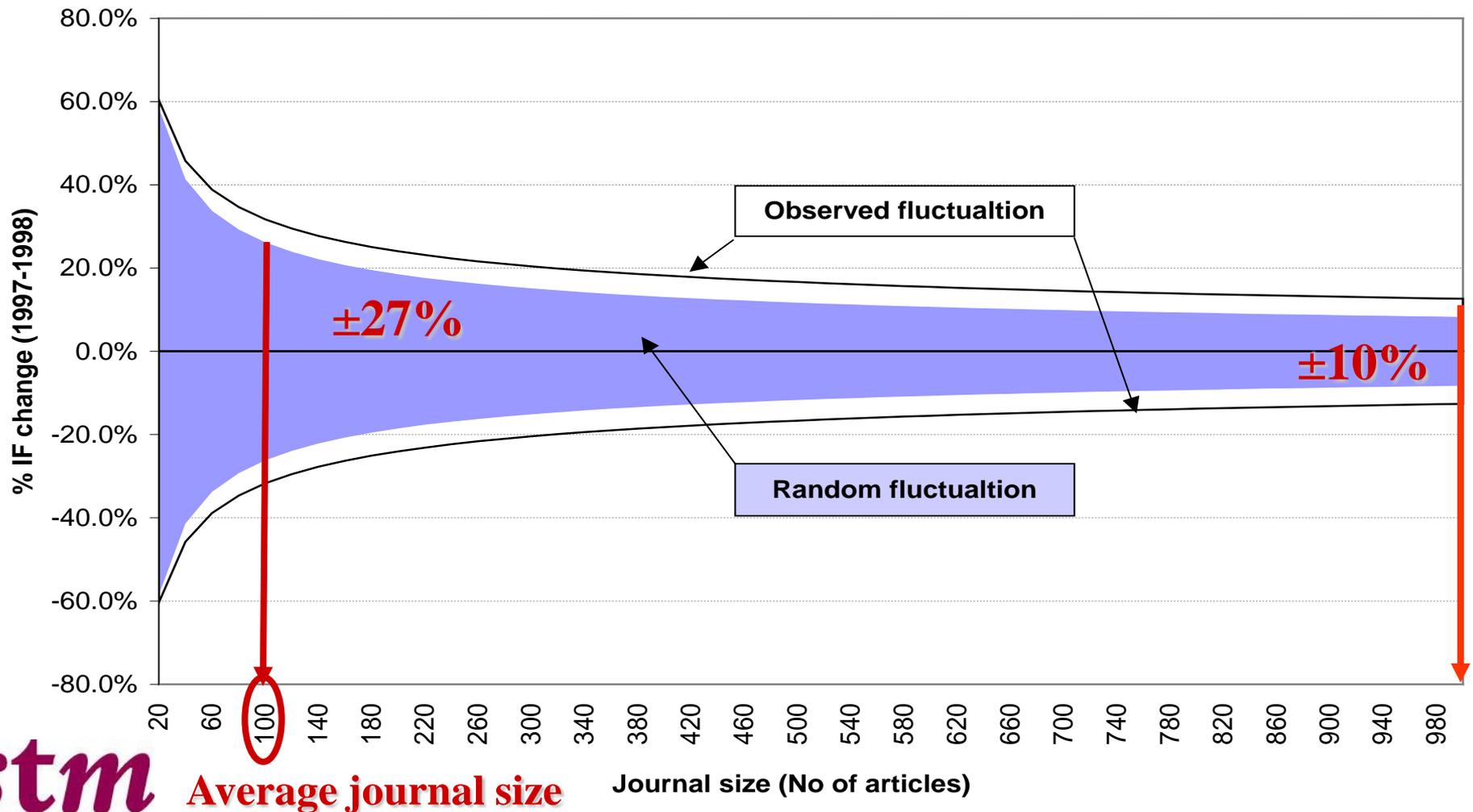


# Influences on Impact Factors: Article Type



# Influences on Impact Factors: Journal Size

## Impact Factor Fluctuation vs Journal Size



# Usage As A Metric

- Assumption: the act of use denotes reading
  - Historical shelving studies by libraries
  - Downloading of electronic full text articles  $\equiv$  reading?
  - Seems that only 50% of downloads get read
- Counting downloads can be misleading
  - Need common (electronic) standards (Project COUNTER) for the count
  - ...and, who's doing the downloading?

# Usage As A Metric

- Need to normalize for comparison
  - journals have different sizes (papers/year) and widely varying potential readerships
  - Very similar problems to citation analysis
- UKSG and publishing community have developed a “journal usage factor”; but won’t help book intensive disciplines

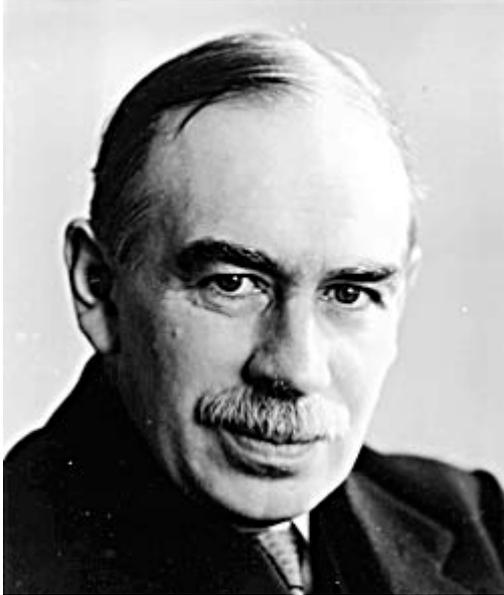
# Altmetrics

- New measures of “impact”
  - Media mentions
  - Newspapers
  - Twitter
  - Policy documents
  - Citations
    - Citeulike
    - Mendeley
  - Software developed
- Article level, web-based indicators

**ISSUES**

***stm***

# John Maynard Keynes on Beauty Contests



It is not a case of choosing those [faces] that, to the best of one's judgment, are really the prettiest, nor even those that average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practice the fourth, fifth and higher degrees.

– J M Keynes *General Theory of Employment Interest and Money*, 1936

# Meta-metric Issues

- Game playing
- “You get what you measure”
  - Institutions re assessment rules
- Impact Factor Engineering
  - Include more review articles
  - Classify low impact content as non citable
  - Include editorials that cite all your articles
  - Publish extremely controversial editorials (medical journals) or articles
- Counting the qualitative
  - Tweets differ in usefulness (“this is crap...” v.?)

# Problems of different metrics

- Qualitative
  - Problems of peer review: human beings!
- Quantitative
  - Normalising for markets and sizes
  - Disaster of numerical values
    - abuse of IFs, dangers of simplistic use of download data
  - Citations are author measures
    - Longer time constants: archival value?
  - Downloads are a reader measure
    - Shorter time constants: sexiness?
  - Article level counts too low for statistical significance?

# Other Evaluation Approaches

- Mainly citation based and author-focussed
  - *h*-index: proposed 2005 by Jorge Hirsch
    - A scientist has index *h* if *h* of [their]  $N_p$  papers have at least *h* citations each, and the other  $(N_p - h)$  papers have at most *h* citations each
      - intended to measure simultaneously the quality and sustainability of scientific output, as well as, to some extent, the diversity of scientific research
    - *g*- index: proposed 2006 Leo Egghe
      - Version of the *h*-index
  - Recursive PageRank algorithm-based
    - Eigenfactor approach
  - PROBLEM: ALL TOO COMPLICATED!!!

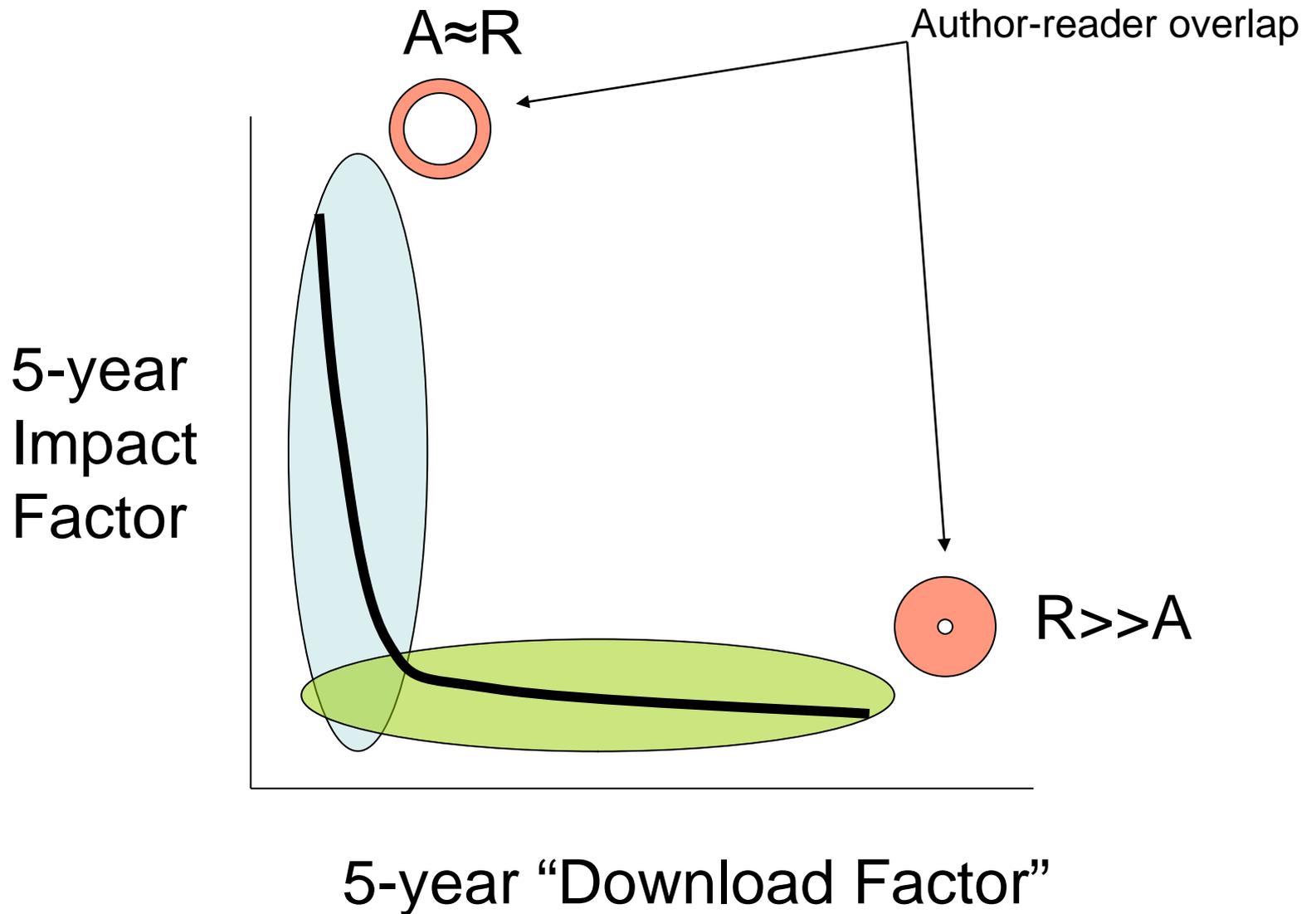
# Many Routes to Salvation...



“Let a hundred flowers bloom, let a hundred schools of thought contend!”

- Mao Tse Tung  
On starting the  
Cultural Revolution  
in 1968

# A 2-D Factor Scatter Plot



# Ways forward?

- Combi-metrics or Oligo-metrics?
  - Multiply sourced metrics giving a rounded flavour from many perspectives
- Group classifications rather than numbers?
  - E.g., like the rating agencies:
  - AAA, AAB+
- The more metrics are seen as flawed, the more academia will cleave to “beauty contests” to which metrics were meant to be the alternative

# Further Reading

- *Asis&t Bulletin* April/May 2013  
[http://www.asis.org/Bulletin/Apr-13/AprMay13\\_Piwowar.html](http://www.asis.org/Bulletin/Apr-13/AprMay13_Piwowar.html)
- CHE 3 June 2013  
<http://chronicle.com/article/Rise-of-Altmetrics-Revives/139557/>
- M Amin & M Mabe Impact factors: use and abuse *Perspectives in Publishing*  
[http://www.elsevier.com/framework\\_editors/pdfs/Perspectives1.pdf](http://www.elsevier.com/framework_editors/pdfs/Perspectives1.pdf)