UNIVERSITY OF HOHENHEIM

Working with scientific literature: Plagiarism and Impact

3502-440 Methods of Scientific Working for Crop Science

WS 2024/2025

Prof. Dr. Karl Schmid
Institute of Plant Breeding, Seed Science and Population Genetics
University of Hohenheim

Outline

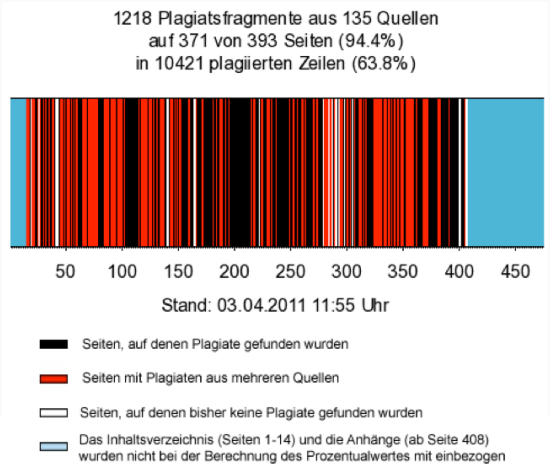
Plagiarism in scientific writing

Searching for scientific literature

Impact and quality of scientific literature
- Criticisms of the impact factor

Examples of plagiarism in Ph.D. thesis

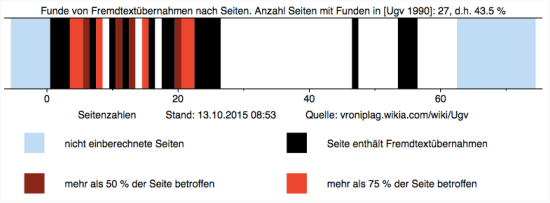
Karl-Theodor zu Guttenberg, Former German Defense Minister



de.guttenplag.wikia.com

Examples of plagiarism in Ph.D. thesis

Ursula von der Leyen, Current Head of EU Commission



de.vroniplag.wikia.com/wiki/Ugv

What is plagiarism?

Plagiarism means using another's work without giving credit.

If you use others' words, you must put them in quotation marks and cite your source.

You must also give citations when using others' ideas, even if you have paraphrased those ideas in your own words.

What is plagiarism?

1. Failure to cite correctly
2. Misconduct: Intent to cheat

⇒ Know the rules of the game!

What are the rules for your studies here in Hohenheim?

- Information on plagiarism are on this link: `https://agrar.uni-hohenheim.de/en/plagiats`
- But you have to sign a form that you do not plagiarize in your term papers, theses and other works.
- Professors can use plagiarism detection software
- There are no common rules regarding the format of citations at the University of Hohenheim
  ⇒ Discuss with your thesis advisor!

Critical issue

What is the thin red line between plagiarism and correct scientific working?

Example of a student's term paper in one of our courses:

and Robertson, 2007). Furthermore, the illegal nature of cultivation in most countries has led to a proliferation of informal names which may not be uniquely or universally applied (Knigt et al., 2010). Medical cannabis-plant breeders use an alternative taxonomy and recognise two taxa, the "species" indica and sativa. In the drug-plant breeding context the taxon indica generally refers to Asian drug plants, especially with features prominent in the Afghani strains, of wide leaflets, compact habit and early maturation (de Meijer & van Soest, 1992). These are strains that have been traditionally used in the production of resin for drug purposes (hashish) as opposed to leaf and inflorescence (flower-spike)

The plagiate was detected with a software.

How does outlining help to avoid plagiarism?

Summarize or paraphrase the ideas, work or thoughts of others

⇓

Outline

⇓

Rewrite with your own words!

Further reading

Since plagiarism is such an abundant problem, you can find easily many books in online bookstores or specific guids on the internet.

Plagiarism in scientific writing

Searching for scientific literature

Impact and quality of scientific literature

Criticisms of the impact factor

How to find literature

- PubMed: `https://www.ncbi.nlm.nih.gov/pmc/` (government-funded)
- Google Scholar: `http://scholar.google.de/` (automatic extraction)
- Scopus: `http://www.scopus.com` (curated, requires subscription, owned by Elsevier)
- Web of Science: `http://webofknowledge.com` (curated, requires subscription, owned by Clarivate Analytics)

How to get the full text of a publication

- Google Scholar link (searches for free online copies)
- University library
- Write to the author - corresponding author
- Use the unpaywall add-on for your browser `http://unpaywall.org`

New developments: Knowledge graphs

Visualization and traversal of citation networks.

Adresses following problem: Relevant literature is missed in a scientific publication

Examples:

- `https://openknowledgemaps.org`
- `https://researchrabbitapp.com`

Why measure the impact of scientific publications?

- Research is expensive → evaluation and ranking
- Measure scientific productivity

Measures of scientific productivity

- Amount of money obtained from external, competitive funds
- The number and quality of scientific publications
- The number of patents

Bibliometric measures of scientific impact

- Bibliometrics: A research method in library and information science
- Basic measure of scientific quality: Impact factor
- The IF was invented by Eugene Garfield (Founder of Institut of Scientific Information, ISI)
- ISI is now part of Thomson Reuters
- Initial definition of IF: Average number of citations of a journal article over time

Formal definition of the impact factor

- A = the number of times articles published in 2006 and 2007 were cited by indexed journals during 2008
- B = the total number of 'citable items' published in 2006 and 2007
- $\Rightarrow$ 2008 impact factor $= A/B$
- 'Citable items' are usually articles, reviews, proceedings, or notes; not editorials or Letters-to-the-Editor.

The impact factor is calculated by the Journal Citation Index of the Thomson Reuters Company.

What is an impact factor?

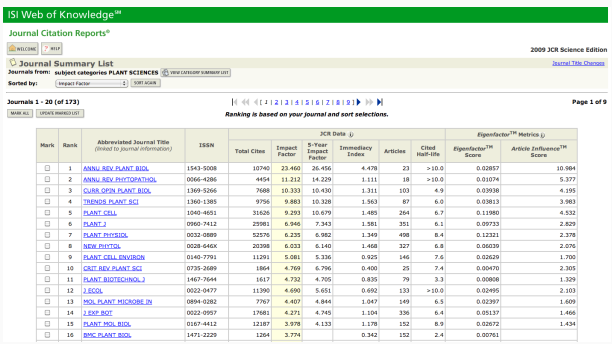
The impact factor ranks journals by their impact.

Example calculation:

- A = the number of times articles published in 2009-2010 were cited in indexed journals during 2011
- B = the number of articles, reviews, proceedings or notes published in 2009-2010
- Impact factor (IF) 2011 = A/B

$$\text{IF} = \frac{\text{number of citations during year to articles from last two years}}{\text{number of articles in the last two years}}$$

Other measures of impact



ISI Web of Knowledge℠

Journal Citation Reports®

2009 JCR Science Edition

Journal Summary List

Journals from: subject categories PLANT SCIENCES

Sorted by: Impact Factor

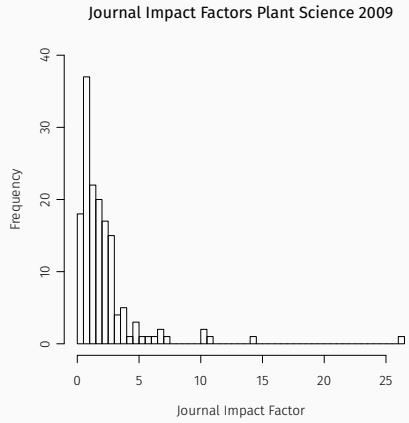
Journals 1 - 20 (of 173)

Page 1 of 9

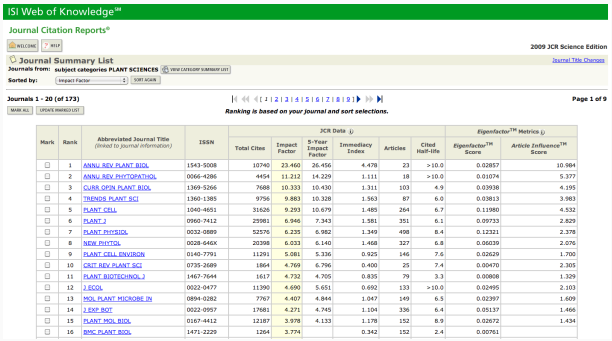
Ranking is based on your journal and sort selections.

Mark	Rank	Abbreviated Journal Title (linked to journal information)	ISSN	JCR Data						Eigenfactor℠ Metrics	
				Total Cites	Impact Factor	5-Year Impact Factor	Immediacy Index	Articles	Cited Half-life	Eigenfactor℠ Score	Article Influence℠ Score
☐	1	ANNU REV PLANT BIOL	1543-5008	10740	23.460	26.456	4.478	23	>10.0	0.02857	10.984
☐	2	ANNU REV PHYTOPATHOL	0066-4286	4454	11.212	14.229	1.111	18	>10.0	0.01074	5.377
☐	3	CURR OPIN PLANT BIOL	1369-5266	7688	10.333	10.430	1.311	103	4.9	0.03938	4.195
☐	4	TRENDS PLANT SCI	1360-1385	9756	9.883	10.328	1.563	87	6.0	0.03813	3.983
☐	5	PLANT CELL	1040-4651	31626	9.293	10.679	1.485	264	6.7	0.11980	4.532
☐	6	PLANT J	0960-7412	25981	6.946	7.343	1.581	351	6.1	0.09733	2.829
☐	7	PLANT PHYSIOL	0032-0889	52576	6.235	6.982	1.349	498	8.4	0.12321	2.378
☐	8	NEW PHYTOL	0028-6400	20398	6.033	6.140	1.468	327	6.8	0.06039	2.076
☐	9	PLANT CELL ENVIRON	0140-7791	11291	5.081	5.336	0.925	146	7.6	0.02629	1.700
☐	10	CRIT REV PLANT SCI	0735-2689	1864	4.769	6.796	0.400	25	7.4	0.00470	2.305
☐	11	PLANT BIOTECHNOL J	1467-7644	1617	4.732	4.705	0.835	79	3.3	0.00808	1.329
☐	12	J ECOL	0022-0477	11390	4.690	5.651	0.692	133	>10.0	0.02495	2.103
☐	13	MOL PLANT MICROBE IN	0894-0282	7767	4.407	4.844	1.047	149	6.5	0.02397	1.609
☐	14	J EXP BOT	0022-0957	17681	4.271	4.745	1.104	336	6.4	0.05137	1.466
☐	15	PLANT MOL BIOL	0167-4412	12187	3.978	4.133	1.178	152	8.9	0.02672	1.434
☐	16	BMC PLANT BIOL	1471-2229	1264	3.774		0.342	152	2.4	0.00761	

Impact factors of Plant Science journals



Other measures of impact



ISI Web of Knowledge℠

Journal Citation Reports®

2009 JCR Science Edition

Journal Summary List

Journals from: subject categories PLANT SCIENCES

Sorted by: Impact Factor

Journals 1 - 20 (of 173)

Page 1 of 9

Ranking is based on your journal and sort selections.

Mark	Rank	Abbreviated Journal Title (linked to journal information)	ISSN	JCR Data						Eigenfactor™ Metrics	
				Total Cites	Impact Factor	5-Year Impact Factor	Immediacy Index	Articles	Cited Half-life	Eigenfactor™ Score	Article Influence™ Score
☐	1	ANNU REV PLANT BIOL	1543-5008	10740	23.460	26.456	4.478	23	>10.0	0.02857	10.984
☐	2	ANNU REV PHYTOPATHOL	0066-4286	4454	11.212	14.229	1.111	18	>10.0	0.01074	5.377
☐	3	CURR OPIN PLANT BIOL	1369-5266	7688	10.333	10.430	1.311	103	4.9	0.03938	4.195
☐	4	TRENDS PLANT SCI	1360-1385	9756	9.883	10.328	1.963	87	6.0	0.03813	3.983
☐	5	PLANT CELL	1040-4651	31626	9.293	10.679	1.485	264	6.7	0.11980	4.532
☐	6	PLANT J	0960-7412	25981	6.946	7.343	1.581	351	6.1	0.09733	2.829
☐	7	PLANT PHYSIOL	0032-0889	52576	6.235	6.982	1.349	498	8.4	0.12321	2.378
☐	8	NEW PHYTOL	0028-646X	20398	6.033	6.140	1.468	327	6.8	0.06039	2.076
☐	9	PLANT CELL ENVIRON	0140-7791	11291	5.081	5.336	0.925	146	7.6	0.02629	1.700
☐	10	CRIT REV PLANT SCI	0735-2689	1864	4.769	6.796	0.400	25	7.4	0.00470	2.305
☐	11	PLANT BIOTECHNOL J	1467-7644	1617	4.732	4.705	0.835	79	3.3	0.00808	1.329
☐	12	J ECOL	0022-0477	11390	4.690	5.651	0.692	133	>10.0	0.02495	2.103
☐	13	MOL PLANT MICROBE IN	0894-0282	7767	4.407	4.844	1.047	149	6.5	0.02397	1.609
☐	14	J EXP BOT	0022-0957	17681	4.271	4.745	1.104	336	6.4	0.05137	1.466
☐	15	PLANT MOL BIOL	0167-4412	12187	3.978	4.133	1.178	152	8.9	0.02672	1.434
☐	16	BMC PLANT BIOL	1471-2229	1264	3.774		0.342	152	2.4	0.00761	

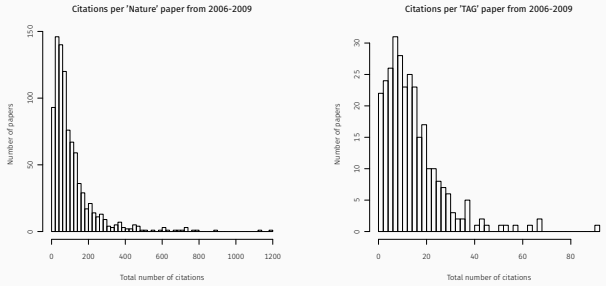
Citations of a general and a specialist science journal

	Impact	Number of	Citations per article			
Journal name	factor	articles	Mean	Median	Min.	Max.
Nature	34.480	905	110.6	71	0	1194
TAG	3.363	298	14.7	12	0	91

Theoretical and Applied Genetics (TAG)

- Chief editor: Prof. A. E. Melchinger, University of Hohenheim
- A leading journal in the field of plant breeding and genetics

Citations of a general and a specialist science journal



Measuring the impact of individual researchers

- Total number of peer-reviewed articles published
- Average impact factor of journals
- Total number of citations of a scientist's papers
- Hirsch-Index (*h* index)
- *b* Index: *h* index / years working in science

What is an h-index?

- Named after its inventor Hirsch Hirsch (2005),`https://de.wikipedia.org/wiki/H-Index`
- Allows the ranking of individual scientists

Definition:

- *n* articles that are cited at least *n*-times
- 1 article 1 times cited: $h = 1$
- 4 articles 4 times cited: $h = 4$

Calculate the *h*-index of the following scientist:

- 1. article, 19 times cited
- 2. article, 8 times cited
- 3. article, 4 times cited
- 4. article, 20 times cited
- 5. article, 10 times cited

The impact of countries or institutions

- Citations per professor
- Euros spent per citation
- Average number of publications per citizen
- Average number of citations per citizen
- etc.

Why are impact factors used?

- Seemingly objective and automated assesment of scientific quality
- Widely applicable (scientists, institutes, journals, etc.)
- An efficient and fair system of reputation building?

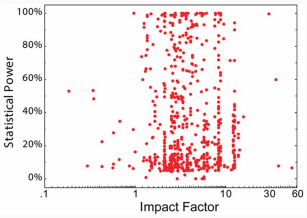
Major criticism of impact factors

- Quantity does not imply quality
- The real value of research becomes often obvious much later
- Many low vs. few high impact papers
- Counting impact factors puts a big administrative burden on scientists.
- Issues of authorship

Validity of the impact factor

- The impact factor is highly discipline-dependent
- The impact factor refers to the *average* number of citations per paper, but this is not a normal distribution ⇒ Pareto distribution (80:20 rule)
- Too many self-citations.

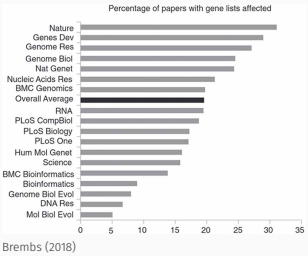
Lack of correlation between impact factor and research quality



Brembs (2018)

No association between statistical power and journal impact factor (IF). The statistical power of 650 eligible neuroscience studies plotted as a function of the IF of the publishing journal. Each red dot denotes a single study

Lack of correlation between impact factor and research quality



Brembs (2018)

Journals with above-average error-rate rank higher than journals with a lower error-rate. Shown is the prevalence of gene name errors in supplementary Excel files as the percentage of publications with supplementary gene lists in Excel files affected by gene name errors.

Manipulation and misuse of the impact factor

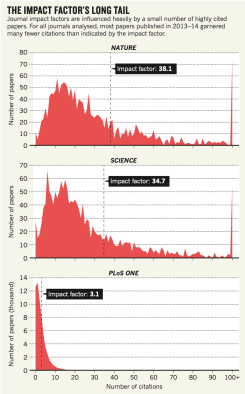
Manipulation

- Journals publish review papers
- Change fraction of 'citable items'
- Ask authors to preferentially cite articles from the same journal.

Misuse

- Non-normal distribution of citations needs to be interpreted correctly
- High-impact journals may have weak articles
- Low-impact journals may have strong articles

Criticism of impact factor by scientists




Publishing elite turns against impact factor

Senior staff at societies and leading journals want to end inappropriate use of the measure.

BY EWEN CALLAWAY

Callaway (2016)

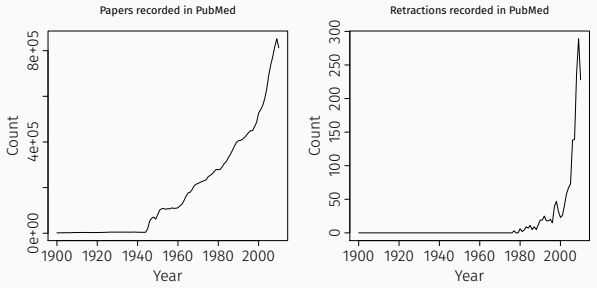
Coping with growing numbers of scientific articles

- Take impact factors with a grain of salt.
- List your five most important papers in applications
- Text mining of papers
- Identify key papers in a scientific field and read them yourself

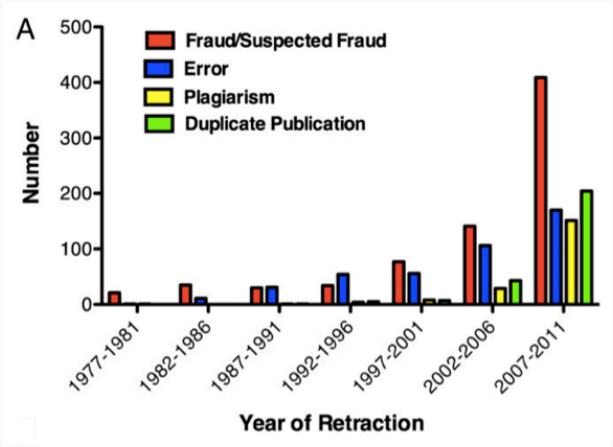
Retractions of papers

- The growing awareness for falsifications.
- New techniques to recognize the manipulation of images.
- A stronger competition between research groups that gives a higher chance that critical and important experiments are being repeated.
- `http://retractionwatch.wordpress.com/`

Retractions of papers

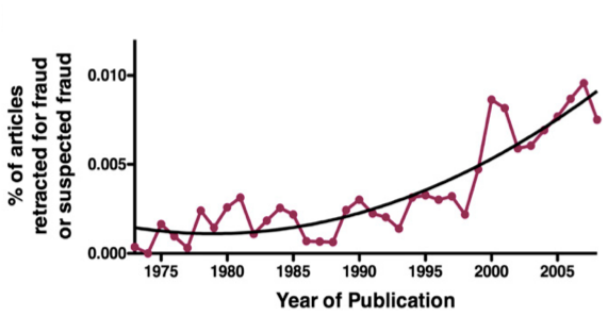


Cause of retractions



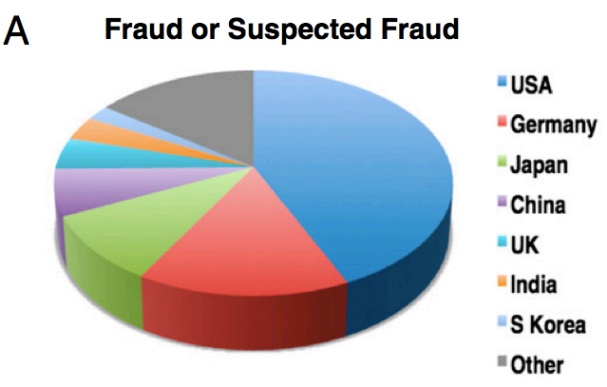
Fang et al. (2012)

Cause of retractions



Fang et al. (2012)

Retraction because of fraud by country



Fang et al. (2012)

References i

Brembs, B. (2018). Prestigious Science Journals Struggle to Reach Even Average Reliability. *Frontiers in Human Neuroscience*, 12.

Callaway, E. (2016). Publishing elite turns against impact factor. *Nature*, 535:210–211.

Fang, F. C., Steen, R. G., and Casadevall, A. (2012). Misconduct accounts for the majority of retracted scientific publications. *Proceedings of the National Academy of Sciences*, 109(42):17028–17033.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46):16569–16572. Publisher: National Academy of Sciences Section: Physical Sciences.