CURRENT RESEARCH IN SOCIAL PSYCHOLOGY

Volume 5, Number 1
Submitted: October 19, 1999
Resubmitted: January 9, 2000
Accepted: January 10, 2000
Publication date: January 20, 2000

THE CITATION IMPACT FACTOR IN SOCIAL PSYCHOLOGY: A BAD STATISTIC THAT ENCOURAGES BAD SCIENCE?

Craig McGarty
Australian National University

[1]
---
[2]

ABSTRACT

The impact factor is widely used to measure journal quality in social psychology even though it actually measures apparent rather than genuine impact. Analyses of citations in two journals show that impact factors may detect some genuine publication impact, but this is a small proportion of apparent impact (32%). Self-citation figures of (61% and 38% of impacting citations) show that authors as often become aware of impact-credited articles by writing them as by reading them. If impact factors are to be used in this field then they should involve at least four years' coverage.

INTRODUCTION

The evaluation of scholarly work often centers around the quality of journals and the need for good measures of this quality. The consensus in social psychology seems to be that journal quality is best measured through citations of papers published in journals in subsequent journal articles.

The opposing views on this point are summarized in the papers by Buffardi and Nichols (1981) and Boor (1982). The former conclude that impact factors are the best measures of journal quality whereas Boor contends that citation-based measures are particularly dubious and subject to abuse. By the time that Rotton, Levitt and Foos's (1993) paper was published this debate seemed to have abated and a citation-based measure had emerged as the preferred measure of journal quality in psychology.

The purpose of this paper is not to enter the debate as to whether citation analysis is valid. Rather I intend to investigate whether the specific measure that is used in social psychology (and in other fields) is a reliable and valid means of detecting genuine publication impact. This is an

interesting question because the research practices of scientists and the publication procedures of their journals vary from field to field. These factors may interact to render a measure appropriate or inappropriate for a particular field. We cannot assume that because a measure may be suitable for internal medicine or particle physics that it is appropriate for our own field.

[2]
---
[3]

In most scientific disciplines and sub-disciplines the formal rankings are based on the ubiquitous Journal Citation Reports provided by the Institute of Scientific Information (ISI) (e.g., from the Social Sciences Citation Index).

Many authors and editors spend much time considering these figures and the newsletters of the Society of Personality and Social Psychology (1997) and the European Association for Experimental Social Psychology (1998) both carry recent analyses of these statistics so that members of these societies can track the progress of their society's flagship journal too (see also Spears', 1998, *British Journal of Social Psychology* editorial in which he notes the rise in that journal's impact but also alludes to doubts about the meaning of impact factors). Impact figures are also used by universities for tenure, appointment, and promotion decisions; granting bodies assessing research funding applications; and by libraries for deciding which journals to stock. It is worth noting that the publications page of the SPSP web-site describes PSPB as "a monthly journal that is now ranked as the second-most cited social psychology journal according to the Social Science Citation Index" (URL http://www.spsp.org/spsp/pubs.htm February 23, 1999). This claim can be challenged on some technical grounds but it does serve to illustrate that impact factors are recognized as a basis for journal rankings in this field.

The supporting literature provided by the ISI (Garfield, 1994) recommends that the impact factor "should be used with informed peer review". One purpose of this paper is to facilitate the development of well-informed views in the field.

The impact factor is based on the number of citations in a given calendar year (by papers published in all journals in that year) of papers published in the two preceding years in a particular journal divided by the total number of papers published in that journal in those two years. Thus the 1998 impact factor for a journal such as Personality and Social Psychology Bulletin (PSPB) involves all of the 1998 citations from all articles in the ISI database of 1996 and 1997 publications in PSPB divided by the number of articles published in PSPB in those two years. This yields a mean impact factor which allows comparisons of journals in the personality and social psychology journal set (which is defined by the ISI). Once these impact factors are obtained the journals are ranked from highest to lowest.

[3]
---
[4]

Most of the journals in the personality and social psychology journal set regularly have mean impact factors of less than 1.0. This implies that according to the measure most articles published in these journals have no impact. It also raises the question of whether the mean should be

selected as the appropriate measure of central tendency under conditions where the distribution is so obviously skewed.

The impact factor has been favored by many because of its objective nature (but see Boor, 1982), but even if objectivity contributes to reliability such high reliability is no substitute for validity. The problems with the validity of the impact factor that I wish to address relate primarily to the two year time in relation to the methods by which authors actually learn about the papers they cite. Papers published in 1998 will tend to be conceived and written at some time earlier than 1998. This necessarily means that some proportion of the research published in 1996 and 1997 was not available to authors in published form at the time they wrote papers published in 1998. The impact factor is thus a misleading statistic because it cannot detect genuine impact where that impact takes much more than two years to express. Reedijk (1998) cites forthcoming research by Moed, van Leeuwen and Reedijk which shows that rankings of journals in chemistry can change dramatically when a time period longer than two years is selected.

I distinguish here between three potential sources of contributions to apparent impact:

1. Author self-citation
2. Late and pre-publication impact
3. Genuine impact

I use the term "author self-citation" to refer to any citation of a paper written by any author of the citing paper.

[4]
---
[5]

Late and pre-publication impact refers to the citation of work that has been published after the citing paper was submitted for publication. If authors cite papers that are published after their own papers are written we can conclude that they became aware of the paper at a late stage of their research, or that they were aware of the paper in preprint form. Both of these possibilities are problematic for assessing journal impact. If the authors have become aware of the paper at a late stage then it is impossible that the paper could have been highly influential in the original conception of the research. However, even where these late citations do reflect important influences it cannot be argued that the cited work has been influential by virtue of its publication in the journal.

One response might be that prepublication impact is an important part of research in this (and other) fields: if the impact factor captures this then that is a good thing. However, the impact factor is actually a poor measure of prepublication impact. Same-year citations and "in press" publications are excluded from impact factors. A January 1999 paper cited in a December 1999 article would not be credited with impact.

The key variable for assessing the impact factor is the publication lag time. The lag I focus on involves the time between receipt of a manuscript at a journal and its publication. This lag places a shutter over the window of opportunity for a paper to reflect the impact of the recently

published literature. If the shutter is too wide then it necessarily follows that impact factors must measure something other than genuine publication impact. This problem is most pronounced where the impact factor is not able to detect instances of genuine impact.

To consider the validity of the impact factor for some typical personality and social psychology journals I conducted an analysis of the 1998 volumes of two journals: Personality and Social Psychology Bulletin and the European Journal of Social Psychology (EJSP). I am not addressing the impact of these two journals but their contributions to impact and their ability to detect potentially genuine publication impact. The logic is if we understand what contributes to impact in this field we will be better able to assess the appropriateness of the impact factor.

[5]
---
[6]

I addressed the following questions:

1. What proportion of the literature is unavailable to authors in published form at the time of manuscript submission? That is, how large is the citation shutter?
2. How much apparent impact is actually author self-citation or late/ prepublication impact?
3. How sensitive is the impact factor to genuine publication impact?

DATA

Data were recorded manually for every article published in 1998 in the Personality and Social Psychology Bulletin (Volume 24) and the European Journal of Social Psychology (Volume 28).

The following data were recorded:

1. Manuscript reception date (averaged to the 15th day of the month).
2. Manuscript acceptance date (averaged to the 15th day of the month). No receipt or acceptance date was given for one EJSP paper.
3. The publication date (assumed to be the 1st day of the month).
4. The number of citations of 1996 journal articles.
5. The number of citations of 1997 journal articles.
6. The number of citations of 1996 and 1997 journal articles written by any author of the citing work (i.e., self-citations).
7. The number of citations of 1996 and 1997 papers that could (generously) have been interpreted to have been published before the citing work was received.

From these data the following variables were calculated:

8. The lag in days between receipt and publication.
9. Citation shutter: the proportion of the literature between January 1996 and December 1997 that was not available to the author in published form at the time they sent their paper to the journal (assuming that the proportion of articles published is constant over any two year period).

RESULTS

The 10th, 25th, 75th and 90th percentiles of the variables are shown in Table 1 for PSPB and in Table 2 for EJSP. Some of the variables have approximately normal distributions so their means and standard deviations are also shown.

One of the interesting features apparent from the tables is that a sizeable proportion of papers make no contribution to impact. In fact, 16% (17 out of 105) of papers in PSPB and 31.7% (19 out of 60) in EJSP have no 1996 or 1997 journal article citations. The standard (two year) impact factor treats these papers as having no attributable intellectual ancestry. It is also worth noting that more than 10% of the papers published in 1998 were received before the 1996-1997 survey period began.

Table 1. Citation and lag statistics for Personality and Social Psychology Bulletin (1998)

	Citations					
	1996	**1997**	**Total**	**Self**	**Late**	**PGPI**
Mean	1.65	0.82	2.47	0.93	0.72	0.81
SD	1.61	1.10	2.03	1.23	1.30	1.14
10%	0.0	0.0	0.0	0.0	0.0	0.0
25%	0.5	0.0	1.0	0.0	0.0	0.0
50%	1.0	1.0	2.0	1.0	0.0	0.0
75%	3.0	1.0	3.5	1.0	1.0	1.0
90%	4.0	2.0	5.4	2.0	2.0	3.0

	Lag Statistics			
	Receipt	**Accept**	**Lag**	**Shutter**
Mean	---	---	695.36	530.47
SD	---	---	162.62	186.52
10%	Oct. 1995	Feb. 1997	504.00	296.40
25%	Apr. 1996	Apr. 1997	594.00	412.00
50%	Aug. 1996	Aug. 1997	655.00	504.00
75%	Nov. 1996	Oct. 1997	777.00	626.00
90%	Mar. 1997	Nov. 1997	929.80	809.00

Notes: 1996 and 1997 citations sum to Total (as do Self, Late and PGPI).
Self= Self-citations.

Late = Citations of articles published after manuscript reception.
PGPI = Potentially Genuine Publication Impact (Total-Self-Late).
Receipt = Month of reception.
Accept = Month of acceptance.
Lag = Time in days between reception and publication.
Shutter = Number of days before Jan 1 1998 that the paper was being processed by the journal.

[8]
---
[9]

Table 2. Citation and lag statistics for European Journal of Social Psychology (1998)

	Citations					
	1996	**1997**	**Total**	**Self**	**Late**	**PGPI**
Mean	1.00	0.38	1.38	0.85	0.10	0.43
SD	1.39	0.64	1.56	1.18	0.35	0.65
10%	0.0	0.0	0.0	0.0	0.0	0.0
25%	0.0	0.0	0.0	0.0	0.0	0.0
50%	1.0	0.0	1.0	0.0	0.0	0.0
75%	1.0	1.0	2.0	1.0	0.0	1.0
90%	3.0	1.0	3.0	2.9	0.0	1.0

	Lag Statistics			
	Receipt	**Accept**	**Lag**	**Shutter**
Mean	---	---	652.95	501.42
SD	---	---	168.45	193.45
10%	Dec. 1995	Nov. 1996	501.00	292.00
25%	Apr. 1996	Jan. 1997	563.00	382.00
50%	Sep. 1996	Jun. 1997	624.00	473.00
75%	Dec. 1996	Oct. 1997	746.00	626.00
90%	Mar. 1997	Dec. 1997	900.00	748.00

[9]
---
[10]

1. What proportion of the cited literature is unavailable at the time of publication of citing works?

The answer to the first question rests on the analysis of the citation shutter. For PSPB the median citation shutter is 504 days and for EJSP it is 473 days. Converting these to percentages: the period unavailable for prior impact for a typical paper is 68.9% and 64.7% of the total period surveyed. In other words, for the typical paper at submission time two thirds of the literature that

could have an impact was yet to be published. Moreover, 90% of papers in both journals had a shutter of more than 290 days. This means at least 40% of the literature was unavailable in published form. This implies that where this unpublished literature was cited it must have been self, late or prepublication impact.

2. How much of the apparent impact of journals is due to self-citation or late/ prepublication impact?

The incidences of self-citation in the two journals are 98 from 259 (37.8% for PSPB) and 51 from 83 (61.4% for EJSP). We can thus conclude that a very large proportion of apparent impact in these journals involves self-citation.

The contribution of late and prepublication impact can be estimated by considering the percentage of citations (excluding self-citations) of papers published after the citing paper was submitted. This was 29% for PSPB (76 out of 259) and 7% (6 out of 83) for EJSP.

The balance of these two figures gives the proportion of citations that cannot be excluded from the category of impact by virtue of journal publication. I term this figure potentially genuine publication impact (PGPI). In PSPB 85 out of 259 (32.8%) impact-credited citations fell into this category (of which approximately 50 were in personality and social psychology journals as defined by the ISI). In EJSP 26 (32%) impact-credited citations could be classed as PGPI. In other words, the maximum plausible figure for publication impact on the original submitted version of the paper in the publishing journal is around a third of the total for both journals. This figure is actually an upper bound nevertheless we cannot rule out the existence of some genuine publication impact. Having said that, the median potentially genuine publication impact per article is 0.0 for both journals.

3. How sensitive is the impact factor to genuine publication impact?

Given the low upper bound for genuine publication impact it is worth considering the ability of the impact factor to detect genuine impact on publications in these journals. To address this matter we can consider the window of opportunity for the average paper to be cited.

Thus, let us imagine a paper published in a social psychology journal on July 1 1996. Excluding impact that the paper may achieve through preprint circulation the window of opportunity for the paper to contribute to impact is January 1997 to December 1998. How long does the reader have to conceive, execute, and write up this research and get it to the journal in order for it to be credited as impact?

The paper's impact will be assessed in 1997 and 1998 and will contribute to the journal's impact factor in those years. Assuming that there is space for the citing paper in the last issue of each journal in 1998 there are about 880 days available for the original paper to have detected impact.

Removing the median receipt-publication lag times of 655 days for PSPB and 624 days for EJSP leaves 225 and 256 days respectively.

Thus, in order for the average paper to be expected to have genuine publication impact with average processing times the citing manuscripts should be delivered to the journal within nine months of the impacting publication date. For long papers in social psychology requiring complex research there is therefore little prospect of genuine significant publication impact being detected in these journals especially for papers published late in the survey period.

DISCUSSION

The results can be easily summarized. For these two journals a large proportion of the literature for which impact is assessed was not available to submitting authors of 1998 papers in published form. A noticeable minority of the papers made no contribution to impact. That is, they cited no 1996 and 1997 journal articles. For the purpose of the two year impact factors these papers simply do not exist.

[11]
---
[12]

The figures show that a very large proportion of 1996 and 1997 citations involved self-citation or late/prepublication impact. Only a minority of citations could be regarded as plausible examples of genuine publication impact. Finally, the journals analyzed appeared to have a very limited ability to detect genuine and significant publication impact with the standard two year impact factor.

One response to these observations might be that the "problem" lies with journal publication lag times and the self-citation practices of authors. This reflects a stark misunderstanding. The impact factor is intended to be a performance measure. If journals in this field have long publication lags, or authors commonly cite their own recent work, then the performance measure should take account of these realities.

Despite the high levels of self-citation there is nothing necessarily wrong with (or unusual about) the two journals selected or their contributors. Self-citation does inflate impact factors but that does not mean that self-citation is, in general, a bad thing. If there were no recent self-citations the task of those who seek to assess the quality of journals would be easier, but the task of readers to make important connections in the literature would be more difficult. Journals are published for the benefit of scholars and not quality assessors.

A key practical detail in considering the prevalence of recent self-citation is that authors can only cite (accurately) recent papers of which they know the full reference details. They are more likely to know these precise details for their own work. I suspect that the relatively long lag times of the journals analyzed actually create the conditions for self-citation to be so highly represented in the impact factors. Furthermore, the differences between the two journals (shown most dramatically in the levels of self-citation and late/ prepublication) may rest in the longer acceptance-publication lag of EJSP (during this lag genuine publication impact would be zero in

most cases). It is also worth noting that self-citation is not more prevalent in the EJSP than PSPB in absolute terms (means of 0.83 and 0.98 self-citations per article respectively). The difference between the journals lies in the fact that self-citation is a much higher proportion of the total number of impact-credited citations for EJSP than PSPB.

[12]
---
[13]

It might be argued that if researchers intended to use a statistic which appeared so indefensible for the purpose of measuring psychological variables we could only hope that they would be quickly disabused of this plan by their critics. That social psychology journals can be ranked on the basis of such a statistic may therefore seem alarming.

Such a reaction may, however, be slightly overstated. Despite the inadequacy of the measure there may be a sizeable correlation between the rankings produced by impact factors and the actual quality of journals. This may be because high impact factors have served to enhance the reputation of certain journals over time. It is possible for a measure to contain systematic bias and still be useful for comparisons or rankings.

Nevertheless there are still large problems with the measure. First, distortion of such impact factors is relatively easy to achieve. The total number of citations that contribute to the impact factors of these journals is not large. A journal such as PSPB that publishes 200 papers in two years needs 300 citations of these papers to obtain an impact factor of 1.5. Given that the average number of references in a paper is around 50 the capacity of a single paper (or set of papers by members of a research group) to have an appreciable affect on these totals is considerable.

Second, given the prevalence of author self-citations there is also a premium for journals to publish the work of authors who cite their own work, or who work in close connection with other research groups. If a paper contains many current self-citations then there may be a good chance that the new paper will in turn be cited by those authors' subsequent (i.e., simultaneous) papers.

I have two pressing concerns that apply even if we rule out the prospect of deliberate distortion by authors or editors. The first is that by using the standard impact factor it is virtually impossible for the published literature to have genuine and significant publication impact on empirical research that involves longitudinal studies, cross-cultural studies, complex experimental field studies, or, new applications for research funding. All of these types of research take more than a year to produce manuscripts. This means that the areas of social psychology that do provide detectable contributions to genuine publication impact will be an extremely biased sub-set of the field. If published research does not inspire research that is relatively quick to do, it will tend to have no (detected) impact at all. Thus, given that we accept that cross-cultural and longitudinal studies are important parts of our field, and their prevalence varies between journals we must conclude that the bias in the measure is problematic. These problems are compounded if the impact factor has the effect as productivity measures can (see Pritchard, 1992) of changing the value and therefore nature of contributions in our field.

[13]
---
[14]

The second concern is that the impact factor is also used (unwisely) to rank journals across the social sciences and within psychology. Unless lag times and citation practices are the same in those other fields comparisons between the fields based on the impact factor will inevitably involve serious distortions.

The two year impact factor is an inappropriately short time period to assess genuine publication impact in personality and social psychology. If impact factors are to be used then they may be more valid if they are calculated over at least a four year period. This would even up perturbations created in single years and also incorporate more genuine impact.

In conclusion, it is worth asking how social psychology found itself to be using a poor measure for assessing a matter that is so important to so many of its practitioners. Haslam and McGarty (1998, in press) have argued that scientific practices can be understood as a process of uncertainty management. In psychology uncertainty is customarily dealt with by measuring statistical uncertainty and reducing methodological. Various other forms of uncertainty are frequently banished from formal consideration in the pages of journals and textbooks. Thus, uncertainty that arises from controversial questions involving political and societal matters which might be embarrassing for the field are frequently swept aside. The impact factor is attractive because its seemingly objective nature and the independent status of the statistic's author (the Institute for Scientific Information) prevents many doubts from ever being formed (thereby banishing uncertainty). The two year impact factor clearly favors journals which publish work by authors who cite their own forthcoming work and who are geographically situated to make their work readily available in preprint form. The measure punishes journals which publish the work of authors who do not have membership of these invisible colleges and is virtually incapable of detecting genuine impact. It is not just a bad measure it is an invitation to do bad science.

REFERENCES

Boor, M. (1982). "The citation impact factor: Another dubious index of journal quality." *American Psychologist*, 37:975-977.

Buffardi, L. C., & Nichols, J. A. (1981). "Citation impact, acceptance rate, and APA journals." *American Psychologist*, 36:1,453-1,456.

European Bulletin of Social Psychology, (1998) 10:2-7. *European Journal of Social Psychology*.

Garfield, E. J. (1994). "The impact factor." *Current Contents*, 29:3-5.

Haslam, S. A. & McGarty, C. (1998). *Doing psychology: An introduction to research methodology and statistics*. London: SAGE.

Haslam, S. A. & McGarty, C. (in press). "One hundred years of certitude: Social psychology, the experimental method, and the management of scientific uncertainty." *British Journal of Social Psychology*.

Pritchard, R. D. (1992). "Organizational productivity." In M. D. Dunnette & L. M. Hough (Eds.) *Handbook of industrial and organizational psychology* (Vol. 3, pp. 443-471). Palo Alto, CA: Consulting Psychologists Press.

Reedijk, J. (1998). "Sense and nonsense of science citation analyses: Comments on the monopoly position of ISI and citation inaccuracies. Risks of possible misuse and biased citation and impact data." *New Journal of Chemistry*, 767-770.

Rotton, J., Levitt, M. J., & Foos, P. (1993). "Citation impact, rejection rates, and journal value." *American Psychologist*, 48:911-912.

Society for Personality and Social Psychology Dialogue, (1997) 12:9. "It's official: we're number two!"

Spears, R. (1998). Editorial. *British Journal of Social Psychology*, 37:387-388.

AUTHOR BIOGRAPHY

Craig McGarty is a Senior Lecturer in Psychology at The Australian National University. He is the author of *Categorization in social psychology* (SAGE, 1999) and (with Alex Haslam) of *Doing Psychology: An introduction to research methodology and statistics* (SAGE, 1998). He is an editor of the *The message of social psychology* (Blackwell, 1997) and of *Stereotypes as explanations: The formation of meaningful beliefs about social groups* (CUP, forthcoming). He may be reached through the mail at the Division of Psychology, The Australian National University, Canberra ACT 0200, Australia, or though e-mail at Craig.McGarty@anu.edu.au.