Taylor & Francis
Taylor & Francis Group

Impact Factors – The Basics- Part 1

Jo Cross,
Market Research Manager, Taylor&Francis

What are Impact Factors?

At the simplest level, journal impact factors give the average number of citations received by articles in a particular journal; essentially, the average number of times that articles in a journal are referenced by other articles. Impact factors were conceptualised by Eugene Garfield and Irving H. Sher in the 1960s as an aid to evaluate journals for inclusion in *Current Contents®* and the *Science Citation Index®*.[1,2]

An average citation measure was needed to eliminate the effects of the size and age of a journal on the total number of citations it receives. Older and larger journals will generally receive more citations because they have larger bodies of previously-published articles available to be cited.

Calculating an impact factor requires a denominator (the total number of articles published) and a numerator (the total number of citations received by those articles). A time period, or 'window,' needs to be defined for both these variables.

The publication window is the period during which the articles included in the calculation were published. The citation window is the period during which citations to these articles were counted.

The term Impact Factor most commonly refers to figures calculated and published by Thomson Scientific each year in the Science and Social Editions of the *Journal Citation Reports (JCR)*. These figures give a two-year impact factor and use very specific publication and citation windows. The citation window here is the Impact Factor year and the publication window refers to the two previous years. Therefore, the 2006 *JCR* Impact Factors (released in 2007) were calculated as follows:

$$\frac{\textbf{Number of citations received in 2006 to articles published in 2004 and 2005 in Journal X}}{\textbf{Number of articles published in 2004 and 2005 in Journal X}}$$

For example, the 2006 Impact Factor for *Systematic Biology* was calculated as follows:

Citations received in 2006 to articles published in *Systematic Biology* in 2004 = 552
Citations received in 2006 to articles published in *Systematic Biology* in 2005 = 308
Total Citations received in 2006 to articles published in 2004 and 2005 = 860

Number of articles published in *Systematic Biology* in 2004 = 59
Number of articles published in *Systematic Biology* in 2005 = 52
Total Number of articles published in 2004 and 2005 = 111

$$\textbf{2006 Impact Factor} = \frac{\textbf{Citations in 2006 to articles published in 2004 and 2005}}{\textbf{Number of articles published in 2004 and 2005}}$$

$$\textbf{2003 Impact Factor for } \textbf{\textit{Systematic Biology}} = \frac{\textbf{860}}{\textbf{111}} = \textbf{7.748}$$

The use of this window has the following implications:

- Each article published is included in the denominator for two Impact Factor years (the two years after the year of publication). Thus, an article published in 2004 will be included in both the

2005 and 2006 Impact Factors.

- Citations received in the year of publication or in the third year after publication or later will not count towards any Impact Factor calculation.

Impact Factors have been adopted for use as measures of quality based on the premise that: *"the value of information is determined by those who use it"*.[3] The idea being that the value of a journal can be measured by the number of times its use is formalised in the form of a citation. Thomson Scientific has suggested that Impact Factors can be useful tools for the management of library journal collections. There are, however, several issues associated with the comparable use of Impact Factor as measures of quality and it is important to understand these before utilising Impact Factor data in any decision making process. In this issue we are going to look in some depth at subject variation in Impact Factors and how the way they are calculated affects this.

Subject Variation in Impact Factors

The average number of citations received by articles during the two years after publication varies considerably across different subject fields. This leads to Impact Factors of very different magnitudes between fields as illustrated in Figure 1. The Impact Factors plotted here are calculated by counting the number of citations in the current year to articles published by all journals in the category in the two previous years then dividing this by the total number of these articles. Thus, these figures represent the average number of times an article in the field has been cited; the Impact Factor of the entire category.

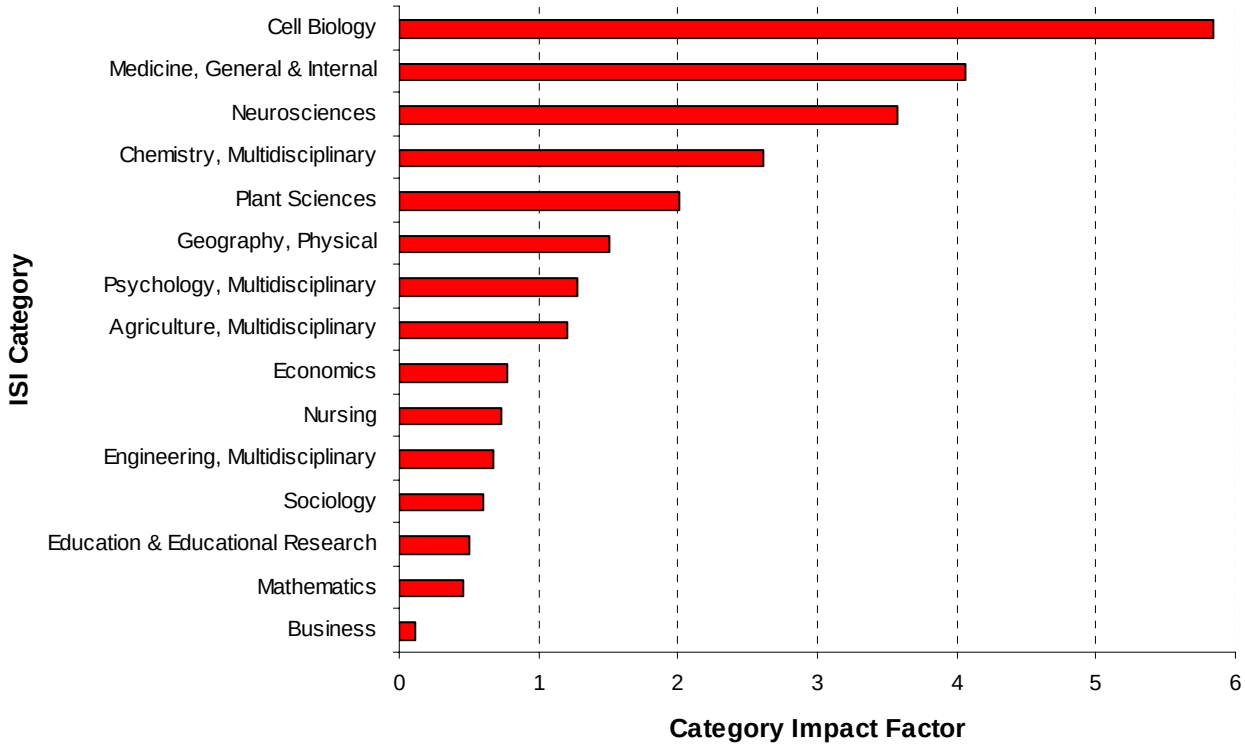

Figure 1: Impact Factors for 14 Categories from the 2003 JCRs

The 14 categories shown in Figure 1 are from a very wide range of fields from both the Science and Social Science editions of the *JCR*. However, even within a more closely-related group of subject

categories, there can be high variation in impact factors. This is illustrated in Figure 2, which shows the category Impact Factors for 10 sub-fields within Physics.

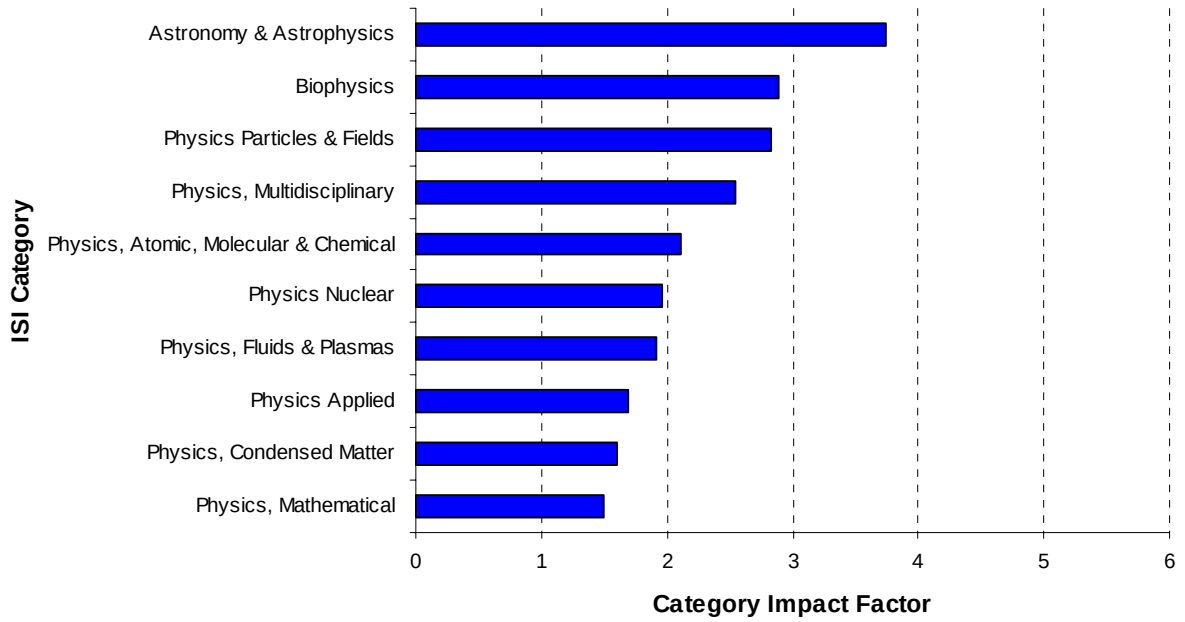

Figure 2: Impact Factors for ten Physics Categories from the 2003 JCR Science Edition

What Causes Subject Variation in Impact Factors?

Subject variations in Impact Factor are due to both different levels and patterns of citation activity to journals in different fields. Figure 3 shows the general magnitude of citation levels after publication for six subject categories. As shown here, there are major differences in the number of citations received by articles in each of the different subject areas with those in Cell Biology, for example, receiving many more citations at all points after publication than those in Economics or Maths.

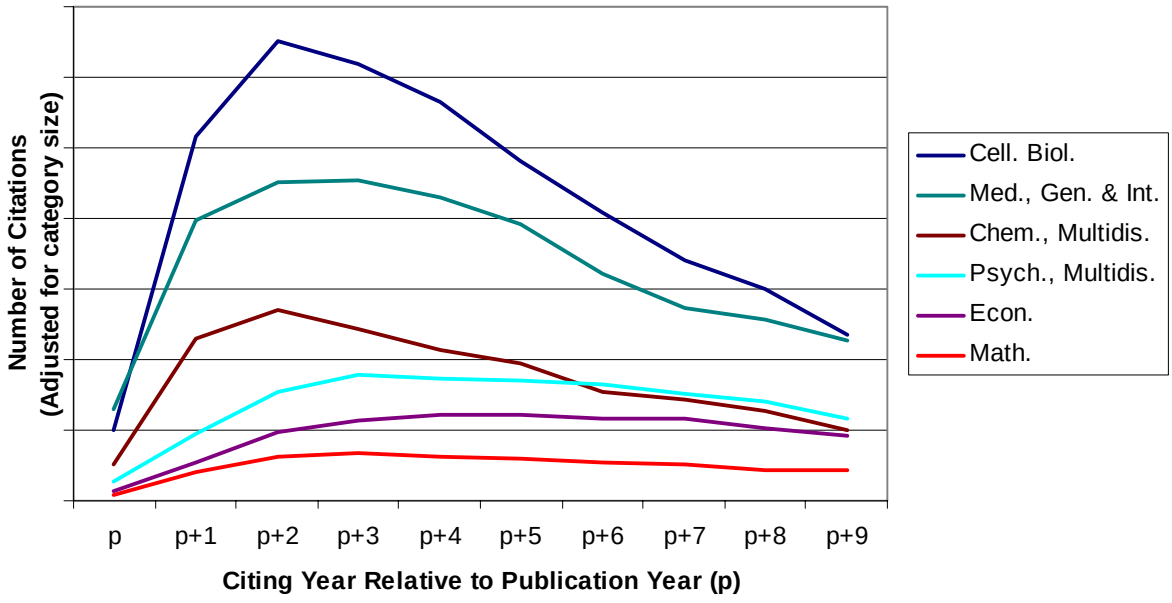

Figure 3: Subject Variation in Citation Levels

One possible reason for higher citation levels in some fields is variation in the average number of authors per article. There is some evidence that highly-cited articles tend to have more authors,[4] and that subject fields with more authors per paper tend to have higher Impact Factors.[5] This is likely to be a consequence of the tendency for authors to cite their own work and that of their research team.

Varying publication behaviours in different subject fields also contribute to differing levels of citations, for example, in many of the Social Sciences greater use is made of books as a method of dissemination than in the hard sciences. This means that a relatively high proportion of references from journal articles in these fields go to books, reducing the total number of inter-journal citations. Equally, many of the citations in these fields may come from references within books and these are mostly not captured by Thomson. In Engineering a similar effect occurs with many citations going to material published in conference proceedings rather than in traditional journals.

Applied fields also differ from basic research fields in terms of the total number of citations received. Journals in applied fields are more likely to reference journals in related basic research fields, than other applied journals. There is no comparable flow of citations back from the basic research journals. Thus, basic research fields tend to receive more citations than related applied fields and, therefore, have higher Impact Factors. For example, journals in basic medicine fields generally have higher Impact Factors than those in clinical medicine fields.[6] We can also see from Figure 2 that the Applied Physics category has a relatively low Impact Factor compared to most of the other Physics categories.

A difference in the level of coverage by Thomson Scientific across subject fields also contributes to varying magnitudes of citation activity as measured by the *JCR*. Citation data in the *JCR* comes only from publications that are indexed by Thomson Scientific. Figure 4, which is based on *JCR* and price list data for 135 commercial and society publishers, illustrates the different levels of coverage of journals by Thomson Scientific across broad fields. Science subjects are in blue and Social Science subjects in yellow.

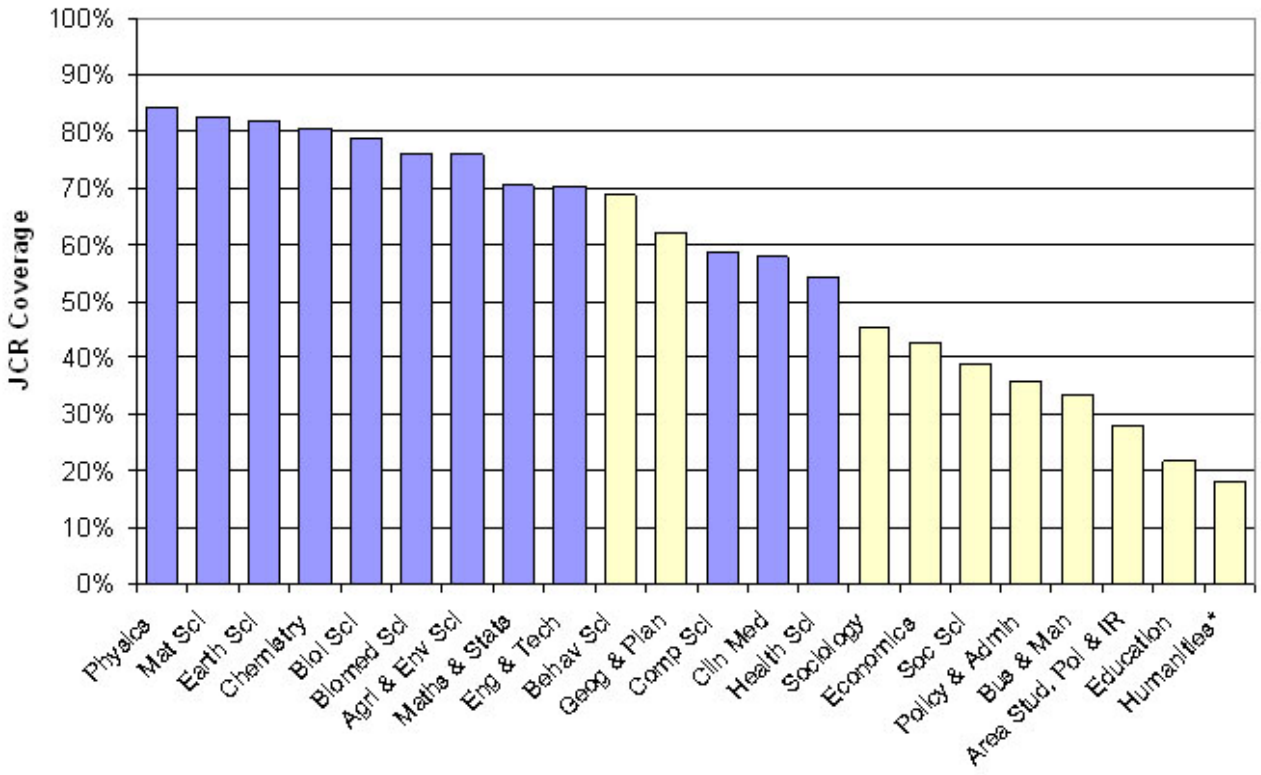

Figure 4: JCR Coverage by Subject – Major Publishers

The source list for citations in the *JCR* is slightly broader than the journals covered within, additionally incorporating any citations from titles covered in the Arts & Humanities Citation Index, ISI Proceedings and the Biosis database. However, these additional sources do not compensate for the different levels of coverage across fields.

In subject fields with lower coverage a higher proportion of citations from the field will be not captured in the *JCR* data set and this will lead to lower recorded citation levels. Figure 4 illustrates generally low coverage in Social Science fields and this is likely to be a contributing factor towards lower recorded citation levels and therefore lower Impact Factors in the Social Sciences *JCR.*

In addition to different levels of recorded citation activity between fields, different citation patterns also affect the magnitude of Impact Factors. Figure 5 shows the percentage of lifetime citations received by year since publication for the same six subject categories. As mentioned earlier, only citations received in the two calendar years after publication count towards an Impact Factor. These citations are those that fall within the Impact Factor window as shown in blue in Figure 5.

Figure 5 clearly illustrates that the proportion of citations falling within the Impact Factor window varies considerably between subject areas. Approximately 22% of life-time citations to Cell Biology articles fall within this window compared to only around 8% of citations to articles in Economics or Maths. This compounds the difference in Impact Factors between these fields since Cell Biology

articles not only receive more citations in total but a higher proportion of these contribute towards Impact Factors.

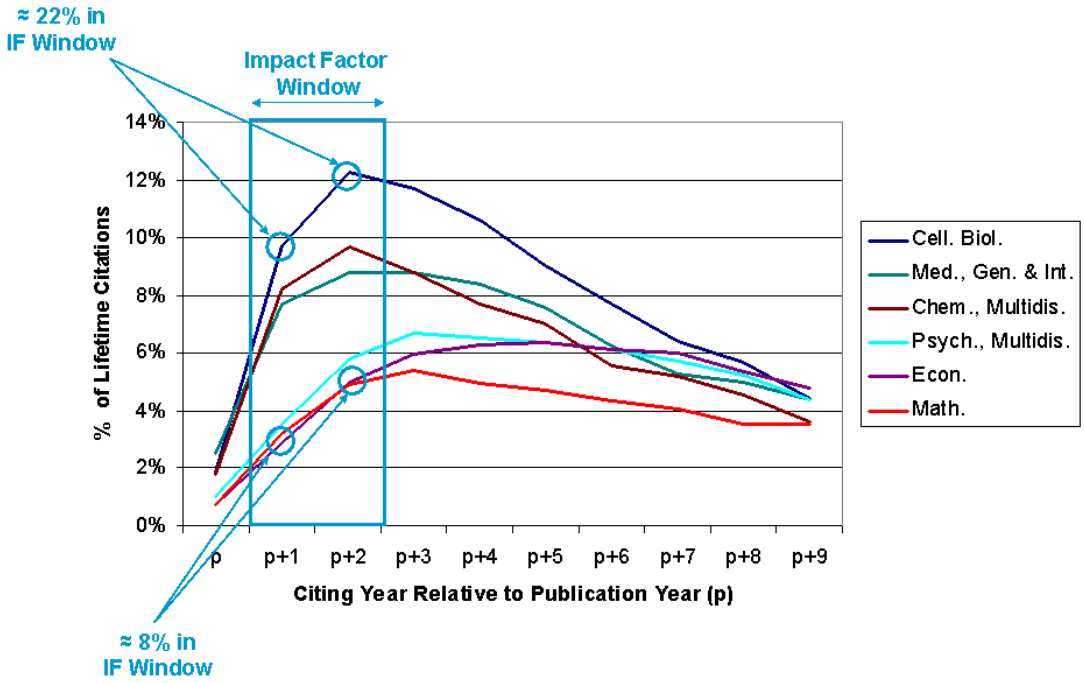

Figure 5: Subject Variation in Citation Distributions

Why a Two-Year Impact Factor?

Impact factors were originally devised to help with journal selection for inclusion in the then ISI's products. At their time of inception, the primary fields of focus for *Current Contents* were fields related to Molecular Biology and Biochemistry.[7] In these fields, 25% of citations received in a particular year were accounted for by articles published in that year and the two previous years.[8] Thus, using a measure that only included citations to recent articles was considered appropriate in this context.

A similar pattern of citations can still be seen in these and other fast-moving fields as illustrated by the pattern of citations (the 'citation curve') for Cell Biology in Figure 5. However, since the citation curve is clearly not the same for all fields many editors have argued for the publication of more long-term impact measures that would take into consideration a greater proportion of life-time citations for slower moving fields.[9,10]

Thomson Scientific justifies its continued publication of two-year impact factors on the basis of two main characteristics of this particular metric: it is both current and responsive.[7]Using citations in the current year* and articles in the two previous years ensures that *JCR* Impact Factors are highly sensitive to recent changes in citation activity.

* 'Current year' here refers to the impact factor year rather than the year of *JCR* publication, which is the year after.

For example, a journal's *JCR* Impact Factor will respond to papers or topical issues that receive an abnormally high number of citations, but will not be skewed for more than two years. Using a longer publication window would dilute these effects but would include them for a longer period. Thomson Scientific considers this sensitivity to be one of the strengths of the *JCR* Impact Factor.

However, it could also be considered a weakness since it means that several years of Impact Factors need to be considered to gauge the general impact of a title.

Another advantage of the *JCR* Impact Factor is that it only requires three years' worth of data to calculate (two publication years and one citation year). Alternative metrics that have been suggested, such as five- or seven-year impact factors, would take over twice as long to produce. Considering the importance that has been, rightly or wrongly, assigned to Impact Factors, this extended delay could be highly detrimental to new journals trying to establish themselves.

Though Thomson continues to stick to the two year measure as the only official Impact Factor published in the annual *JCR,* it does acknowledge that a longer term impact measure may be more appropriate in some fields and suggests a couple of ways in which a five-year measure can be calculated from data available in the *JCR*[11,12]. Extended impact measures for journals can also be calculated from data available in another Thomson Scientific product, the *Journal Performance Indicators*[13]. This product allows comparison to field-specific base lines.

Consequences of Subject Variations in Impact Factors

The high variation in average impact across different subject areas means that Impact Factors can not be used to compare journals from different subject areas. This is why the journals covered by the *JCR* are classified into fairly narrow subject categories. It is only at this level that journals are ranked according to their Impact Factors.

Even within the *JCR* subject categories some journals will have a subject advantage over others, for example, where a subject category contains both basic research and applied journals - practice-based and educational journals often have particularly low impact factors compared to the basic research journals in listed in the same categories. There can also be a subject advantage in categories which are multidisciplinary by nature, for example, social science journals such as the *Journal of Sports Management* and *Journal of the Philosophy of Sport* in the Sports Sciences categories, tend to rank lower than the majority of the sports medicine titles in the same category. All of these things need to be considered when comparing journals on the basis of rank within a subject category.

Subject variations in Impact Factors are also one of the reasons why it is unwise to create an 'average Impact Factor' for a publisher's entire list of journals. Now that many publishers offer 'bulk sales deals' it can be tempting to try and gauge the quality of such an offering by creating an average Impact Factor for the journals included in the deal. However, if subject differences mean that the Impact Factors of journals from different fields can't be compared, then it follows that the data shouldn't be combined either. This sort of publisher-wide average impact factor will always favour publishers with strong Life Sciences programs over those with stronger Social Science programs.

Since Impact Factors can vary even between quite closely related fields as shown in Figure 2, averaging Impact Factors across even a subject package is unlikely to give a fair view of the comparable quality of the package. The same is true of trying to compare publishers on the basis of price per Impact Factor where journals from multiple *JCR* categories are included in the analysis, for example, the LISU report on Trends in Scholarly Journal publishing which combines such diverse categories as Educational Psychology and Cell Biology into one analysis on Biomedical titles[14].

That's it for now on subject variations in Impact Factors. In the next newsletter we will look at some of the other important issues surrounding the calculation and use of Impact Factors.

---

[1] 'The Impact Factor', Thomson Scientific Essay - Originally published in the Current Contents print editions June 20, 1994.
http://scientific.thomson.com/knowtrend/essays/journalcitationreports/impactfactor/
[2] Garfield, E. (1998) *Long Term Vs. Short-Term Journal Impact: Does It Matter?* The Scientist, Vol 2(3): 10-12.
http://www.garfield.library.upenn.edu/commentaries/tsv12(03)p10y19980202.pdf
[3] History of Citation Indexing:
http://scientific.thomson.com/free/essays/citationindexing/history/
[4] Aksnes, D.W. (2003) *Characteristics of highly cited papers*. Research Evaluation, Vol 12(3): 159-170.
[5] Amin, M. & Mabe. M. (2000) *Impact Factors: Use and Abuse*. Perspectives in Publishing, No.1. Available at:
http://www.elsevier.com/framework_editors/pdfs/Perspectives1.pdf
[6] Seglen, P.O. (1997) *Why the impact factor of journals should not be used for evaluating research*. BMJ Vol 314:497.
http://bmj.bmjjournals.com/cgi/content/full/314/7079/497
[7] Garfield, E. (1998) 'The use of journal impact factors and citation analysis for evaluation of science'. Paper presented at the 41st Annual Meeting of the Council of Biology Editors, Salt Lake City, UT, May 4, 1998 - April 17, 1998.
http://www.garfield.library.upenn.edu/papers/eval_of_science_CBE(Utah).html
[8] Garfield, E. (2003) *The Meaning of the Impact Factor*. International Journal of Clinical and Health Psychology, Vol 3(2): 363-369.
http://www.garfield.library.upenn.edu/papers/meaningofif2003.pdf
[9] Garfield, E. (1998) *Long Term Vs. Short-Term,* The Scientist, Vol 2(3): 10-12
[10] Garfield, E. (1998) 'The use of journal impact factors', Salt Lake City, UT
[11] *How to Calculate a Five-Year Impact Factor- JCR* Online Help Section:
http://admin-apps.isiknowledge.com/JCR/help/h_fiveyr_if.htm
[12] 'The Impact Factor', Thomson Scientific Essay - Originally published in the Current Contents print editions June 20, 1994.
[13] Pringle, J. (2008) *Trends in the use of ISI citation databases for evaluation*. Learned Publishing, 21 85-91,

http://dx.doi.org/10.1087/095315108X288901

[14] White, S. & Creaser, C. (2007) *Trends in Scholarly Journal Prices 2000-2006*, LISU Occasional Paper No. 37,
http://www.lboro.ac.uk/departments/ls/lisu/pages/publications/oup2.html