Viewpoint

Edited by A Syntetos

Research contribution: Citation and content analysis

Mingers provides an interesting introduction to and critique of various citation indices and their application (Mingers, 2009). This was my first introduction to the h-index, and it seems to be intuitively attractive. League tables and the like always make fascinating reading even when they come with some warnings of their limitations. Clearly in the future I will have to watch my h-index as well as my BMI, PSA, cholesterol, blood pressure, and driver penalty points. I was surprised at how well I appeared to be doing so I decided to look up my results and this led me to the following observations and conclusions.

Using the Publish or Perish (PoP) software as suggested by Mingers was not difficult in my case as I have a relatively short list of publication, my name is not common and I have two initials. I searched using 'R Ormerod' OR 'RJ Ormerod' OR 'Richard Ormerod'. The search did result in a few papers that were not mine but these could be easily excluded. PoP uses the Google automated system to make a very wide trawl for citations; as a result the number of citations was found to be 483 in my case. As Mingers warned, there was a large tail of cited papers, which the Google software had not managed to marry up with other citations of the same paper; as a result the citations of those papers that affected the h-index were reduced. However, this would be the same for all authors and there is no reason to suppose my case is unusual. In Mingers' paper my h-index was 10 based on 2007 data. Time has moved on and it is now (August 2009) shown by the PoP software to be 12. I also discovered that for some reason one of my papers was missing in the PoP analysis; this paper has 14 citations according to the Web of Science (WoS) and 15 according to Scopus.

I suspected that my relatively good performance (given my short career as an academic) was because I often cited my own papers, particularly when writing *JORS* Viewpoints. The easiest way to examine this seemed to be to use the Scopus citation database where self-citation can be excluded by ticking a box. I therefore searched the Scopus database. Without excluding self-citation my h-index according to Scopus (also in August 2009) was 11. The two sets of papers determining the scores by PoP and Scopus (in other words the top 11 or 12 papers)

were not identical but 10 papers were included in both lists. The fact that the Scopus h-index is one point lower that the PoP/Google h-index is understandable given the more inclusive approach taken by the Google search for citations; in my case, relative to Scopus, the wider trawl outweighed the affect of missed attribution of citations noted above. Scopus identifies 291 citations in total. When the 'exclude self-citation' box is ticked the citations drop to 192 and my h-index to 9. Clearly I am a serial self-citation offender. On the whole, I do not think that citing one of my own papers can be taken to indicate anything about the paper's quality and impact; it hardly provides independent verification and it is open to manipulation (by me). Others might also be offenders but including self-citation seems unfair to those that do not exhibit this trait. In my view, if researchers are to be compared using citation analysis, I think it would be better to omit self-citation.

The 291 citations (including self-citations) found by Scopus contrasted with the 483 found by Google. On closer examination the Google citations included such dubious sources as blogs. Google (and therefore PoP) seems to be too inclusive unless the purpose is to conduct a casual popularity vote. Scopus seemed to include the papers I would have expected to feature and the citations seemed solid. I also tried the WoS. The results here were similar to Scopus, some papers doing better, some worse. Without exclusion of self-citations the WoS h-index was 10. However, there were a number of papers with precisely 10 citations, so it seems likely that before too long the h-index will increase to 11 and the result will be the same as Scopus. Examining the papers below the h-index bar in both Scopus and the WoS suggests that it could be some time before an h-index of 12 is achieved. Google/PoP, which already gives a score of 12, thus seems in my case to inflate the score a little relative to Scopus and WoS.

A further issue is that of multiple authors. In my Google/PoP h-index there are two papers to which I contributed less than 25% and one with a contribution of about 50%. A deduction of about two from my h-index would therefore seem reasonable; this results in a PoP h-index of 10. In the case of Scopus, when self-citations are excluded, there is only one joint paper within my h-index papers; my contribution to that paper was about 50%; this suggests that my h-index is only 8.5 (based on Scopus, excluding self-citations and adjusting for multiple authorship). If the adjustment is to be done automatically a crude adjustment may be necessary such as assuming that each author contributes equally. The h-index for authors who frequently publish with others

will be much more severely affected by such adjustments. Whether the method I have used to adjust my h-index is appropriate or not, clearly some sort of adjustment is needed if comparisons are to be made.

The choice of citation database obviously depends on what one is trying to achieve and what the various databases offer both in terms of the citations accessed and ease of analysis. The title of Mingers' paper refers to 'research contribution' and in the body of the paper a variety of uses is suggested including informing hiring and promotion of individuals, funding, or continued existence for groups and the relative standing of journals. Later the paper examines 'productivity' in terms of the h-rate index and illustrates how a graph against time can be considered to show a researcher's productivity in relation to a stochastic frontier or can be considered output efficient in data envelopment analysis (DEA) terms. However, DEA as I understand it, generally pays a good deal of attention to the nature of inputs and outputs. Following this line of thinking, it would be highly desirable to measure the impact of a paper in terms of its contribution to practice rather than to other theoretical papers; but there are practical difficulties in doing this. One approach could be to examine exclusively the citations contained in application case studies. However, the practical difficulties of conducting such an analysis are formidable as there have been very few genuine case studies of OR interventions published. (The net could be widened somewhat if the qualifier 'genuine' were to be relaxed.)

The fundamental question is what do we mean by research contribution? In practice we cannot measure contribution to practice, so we fall back on measuring contribution to further research via citations. This tends to draw attention away from the purpose of OR academic research and may well encourage irrelevant research. For example, I have had a paper published in *JORS*, which advocates pragmatism. It has caught the attention of a few researchers who are now citing it. No doubt I will find opportunities to cite their understanding of the ideas of pragmatism. Already citations are accumulating. No doubt in time others will want to compare pragmatism with their preferred approach; yet others may wish to attack it as wrong-headed. The result will be plenty of papers, multiple citations, and enhanced h-indices but possibly no discernable contribution to practice. In time the ideas generated could have some impact; but equally the debate could just move on with the paper and its contents forgotten. It is, of course, a contribution to academic debate (as the citations indicate) and deserves some recognition but it is of a very different character from an innovation (say an improved algorithm) that is tested and implemented in an organisational setting. Another way of understanding research output is therefore to concentrate on the content of the research presented in papers (Ormerod, 1997, 2000).

Although citation data is easily accessible and is regularly updated, evaluating content in some way involves someone reading (at least part of) the papers. Content analysis has

| | The Objective of Research | | |
	(gain) Knowledge	(perform) Function	(determine) Purpose
Material	62	30	0
Social	9	20	8
Personal	0	1	0

The Object of Research

Figure 1   The object-objective matrix used to classify the outputs of journal papers (Ormerod, 2002).

not therefore been adopted and the methods have yet to be established. One possible approach has been developed by Reisman and Kirschnick (1994, 1995) and applied to the evaluation of flagship US management science/operational research (MS/OR) journals. These methods were confirmed and the analysis extended to European journals by Ormerod and Kiossis (1997) and Ormerod *et al* (2000). Another possible classification of content uses what I refer to as the objective-objective matrix (Ormerod and Kiossis, 1997); an example is provided in Figure 1 showing an analysis of the papers in the 2000 volume of *JORS*. Each paper was assessed as to whether the object of the research is material, social, or personal, and whether the object is to gain knowledge, perform a function, or determine the purpose of some activity. The paper was then placed in the appropriate cell in Figure 1. Such an analysis required a good deal of reading, but in terms of journal evaluation the requirement for an analysis is infrequent and the results fairly stable from one year to the next. The work required to apply content analysis (of whatever flavour) to groups of researchers could be made considerably easier by using the h-index to limit the number publications considered for each member of the group. Thus a three-step procedure could be envisaged:

1. identify those papers that contribute to each member of the group's h-index;
2. classify the content of each paper identified in step one according some chosen schema (such as the object-objective matrix);
3. interpret the results in some appropriate way.

Thus a group of 10 researchers might have an average h-index of 5 requiring 50 papers to be classified. As a comparison, the analysis shown in Figure 1 involved classifying 132 papers. A group could be identified as belonging to a particular organisational unit (for instance, the Warwick MSOR group) or by researches in different organisations sharing a common interest in a particular field of research (for instance simulation or some subset of simulation).

I see no reason why the research output of individual researchers should not be similarly understood in terms of some classification of the content of their output. What interpretation should be put on such a characterisation is a different matter. My own interpretation of the object-objective matrix shown in Figure 1 uses the standard distinction between research, development, and application: papers in the first column are theoretical in nature and can be classified as research; in the material-function cell the contribution can be considered as development; those in the social-function and social-purpose cells report applications. The classification would allow comparison in DEA terms of researchers who excel in a particular type of output, say theoretical research, with those with a similar emphasis. In summary, although the use of citations for comparative purposes is well established, content analysis provides an opportunity to develop more relevant measures. It is also a challenge.

My conclusion is that the h-index is a useful extension of citation analysis; it provides a summary measure with properties that are intuitively attractive, easy to understand, and is readily obtainable with the tools described by Mingers. My preference is to exclude self-citation and to adjust the overall score for multiple-authorship: much as I enjoy the idea that my h-index is at this time a laudable 12 according to PoP/Google, I think the more modest figure of 8.5 (based on Scopus, excluding self-citations and adjusted for multiple authorship) better reflects my contribution. Better still would be an analysis of the content of my research. Analysis of content, where practical, would be my preferred approach. The h-index can be used to reduce the work required to analyse the output of a group of researchers in terms of the content of their output and thus makes such an analysis more practical.

References

Mingers J (2009). Measuring the research contribution of management academics using the Hirsch-index. *J Opl Res Soc* **60**: 1143–1153.

Ormerod RJ (1997). An observation on publication habits based on the analysis of MS/OR Journals. *Omega* **25**: 599–603.

Ormerod RJ (2000). Is content analysis either practical or desirable for research evaluation? *Omega* **28**: 241–245.

Ormerod RJ (2002). On the nature of OR: Taking stock. *J Opl Res Soc* **53**: 475–491.

Ormerod RJ and Kiossis I (1997). OR/MS publications: Extension of the analysis of US flagship journals to the United Kingdom. *Opns Res* **45**: 179–187.

Ormerod RJ, Delabassi P and Morris C (2000). Research strategies used by OR/MS workers: Extension of the analysis of US flagship journals to the United Kingdom. *Opns Res* **48**: 527–534.

Reisman A and Kirschnick F (1994). The devolution of OR/MS: Implications from a statistical content analysis of papers in flagship journals. *Opns Res* **42**: 577–588.

Reisman A and Kirschnick F (1995). Research strategies used by OR/MS workers as shown by an analysis of papers in flagship journals. *Opns Res* **43**: 731–740.

RJ Ormerod
Warwick University

Editor's Note: The author of the paper discussed in this Viewpoint did not wish to make a further comment.