2-2008

Mapping the Discipline of the Olympic Games an Author-Cocitation Analysis

Peter WARNING
Temasek Polytechnic

Rosie CHING
Singapore Management University, rosieching@smu.edu.sg

Kristine TOOHEY
University of Technology, Australia

SMU
SINGAPORE MANAGEMENT
UNIVERSITY

Mapping the Discipline of the Olympic Games
An Author-Cocitation Analysis

Peter Warning, Rosie Ching Ju Mae, Kristine Toohey
February 2008

Mapping the Discipline of the Olympic Games
An Author-Cocitation Analysis
S33

Peter Warning, Ching Ju Mae Rosie
Temasek Polytechnic, Singapore
Kristine Toohey
Program Manager, Communication Services
Sydney Organising Committee for the Olympic Games,
University of Technology, Australia

The authors conducted an author cocitation analysis on prominent authors writing about the Olympics during the 1990s. Author cocitation is an established bibliometric technique that can be used to measure the relative similarities of topics written about by the cited authors. This enables a visual representation of the "intellectual space" of the discipline, in this case the Olympics, to be created for the period under review. So core and peripheral research areas are identified, along with their major contributors. The representation appears as a two-dimensional cluster-enhanced map. Subject expertise was then applied to the results to place labels on the generated clusters of authors and their topics.

1. Introduction

When most people think about the Olympic Games it is usually in terms of athletic performance. Clearly they are more than that (see Toohey & Veal, 1990). Even the mass media does not confine itself to covering only the sporting angle. For example, symbolism, economic factors, nationalism and politics routinely appear in mass media articles relating to the Olympic Games. There are scholarly journals that are devoted exclusively to the Olympic Games, such as *Olympika* and the *Journal of Olympic History*. So what do we mean when we talk about Olympic scholarship? Cursory scanning of other sport journals also reveals a plethora of subjects ranging from legal aspects to history to philatelic aspects among a host of Olympic topics. This paper questions how can we identify, classify and measure them.

2. Author cocitation analysis

An established method for identifying the different components of a discipline is author cocitation analysis (ACA). ACA is a bibiometric technique that enables a map of the discipline, over a finite time period, to represented. ACA was pioneered by, among others, White and Griffith (1981) from the Drexel College of Information Studies, and Small (1973) and Garfield from ISI, who produce the computerised citation databases *SciSearch* and *Social SciSearch*. Since these early developments, the techniques have been applied to an expanding range of disciplines. McCain (1990) describes a variety of cocitation techniques. The authors have previously applied ACA to physical education pedagogy (Toohey & Warning, 1994).

When applying ACA, we are measuring the number of times that two authors have been cited together in another paper. The premise is that if two authors are cited together, then there is a topical relationship between them. Where there is a group of authors who are being measured, ACA will measure relative strengths of topical relationship between the individual authors within the group. The graphical representation or map of the discipline resulting from ACA and multi-dimensional scaling will assign unique coordinates to the authors. The authors will be clustered together in various groups;

some will be located centrally; and others will be on the periphery, depending on their number and pattern of cocitation with the other authors. Clusters and individuals can be labelled, enabling the various schools of thought and "invisible colleges" within the discipline to be identified. Subject or field knowledge, in assigning labels and relationships, supplements the computer results.

3. Methodology

The first step in conducting ACA is to identify a list of leading and representative writers in the discipline over the targeted time period. There are a number of different techniques for generating this list. For this study, a preliminary literature search was undertaken on the *SportDiscus* database, using subject terms designed to generate articles relating to the Olympic Games. This resulted in thousands of articles. For ACA, articles that are not cited are useless. So certain types of article were discarded, including magazine (such as *Sports Illustrated*) and newspaper (such as the *Weekend Australian*) articles, and articles from all publications less than four pages long. The articles were then sorted by author to identify the most prolific. A list was generated of authors with three or more articles published during the period. This list was supplemented by authors who were known to be influential in the area, but not necessarily prolific writers. There were 56 authors in this list.

Volume of authorship is not a true reflection of importance. One of the greatest influences that an author has is represented by the number of times he or she is cited by others. Using the *Social SciSearch* database, gross citation figures were then identified for each of the authors. The weakness with this approach is that authors with relatively common names, the "Smiths and Jones" of our world, will throw up inflated citation figures as they include other authors with the same family name and first initial. This is a problem only at this stage, as when they are matched with other authors during the cocitation phase, we are left with the "Smiths or Jones" that are the writers within the target discipline. After the citation check, the list was reduced to 37 authors.

Cocitation counts were generated for these authors using the simple command on Social SciSearch:-
s ca=kidd b? and ca= macaloon j? (for authors Kidd and MacAloon)

This search was repeated until every author had been matched with every other author. The results were placed into a square matrix.

For the data to be statistically processed the matrix has to be complete, that is there have to be two entries for each combination of authors. Also the diagonal, representing the same author in both column and row must be filled. Following the methodology of White and Griffith (1981a) it is filled with the highest level of cocitation that they have with another author on the list. So for example, Kidd was cited 19 times with Hargreaves, more than with any other author, so "19" was inserted into the matrix in the Kidd diagonal. A section of the matrix is depicted in table 1.

Table 1
Cocitation matrix extract

	Kidd B	MacAloon J	Hoberman J	Lenskyj H
Kidd B	19	7	9	10
MacAloon J	7	19	17	0
Hoberman J	9	17	17	3
Lenskyj H	10	0	3	26

Authors who had been cocited with less than three of the other authors were then eliminated from the final group, as they would otherwise skew the statistical computations. This left 24 authors in the final author list.

The data in the completed matrix was converted to Pearson product-moment correlations to reduce scale effects. These data were then input into principle components analysis. Cluster analysis was applied to the resulting factor scores. For the non-statistically minded, this process may sound complicated but it can be done with a statistical software package such as SPSS. The ultimate result of all of this is a two-dimensional map (figure 1), plotting the authors according to their cocitation patterns with the other authors on the map.

The authors can then be divided to a range clusters according to their cocitation patterns. With our list of 24 authors there are potentially between 2 and 23 clusters. These are commonly represented in ACA studies by icicle plot or dendrogram. These show, in their different ways, the successive grouping of individual authors and clusters from the situation where all are separate to the situation where they are all joined within one cluster. There is no ideal cluster number for any group of authors. Identification of clusters does not preclude analysis of subgroups within each of the clusters nor how the clusters may then link with each other. Field knowledge of the subject area assists in determining where to divide the clusters.

Figure 1
Author cocitation map
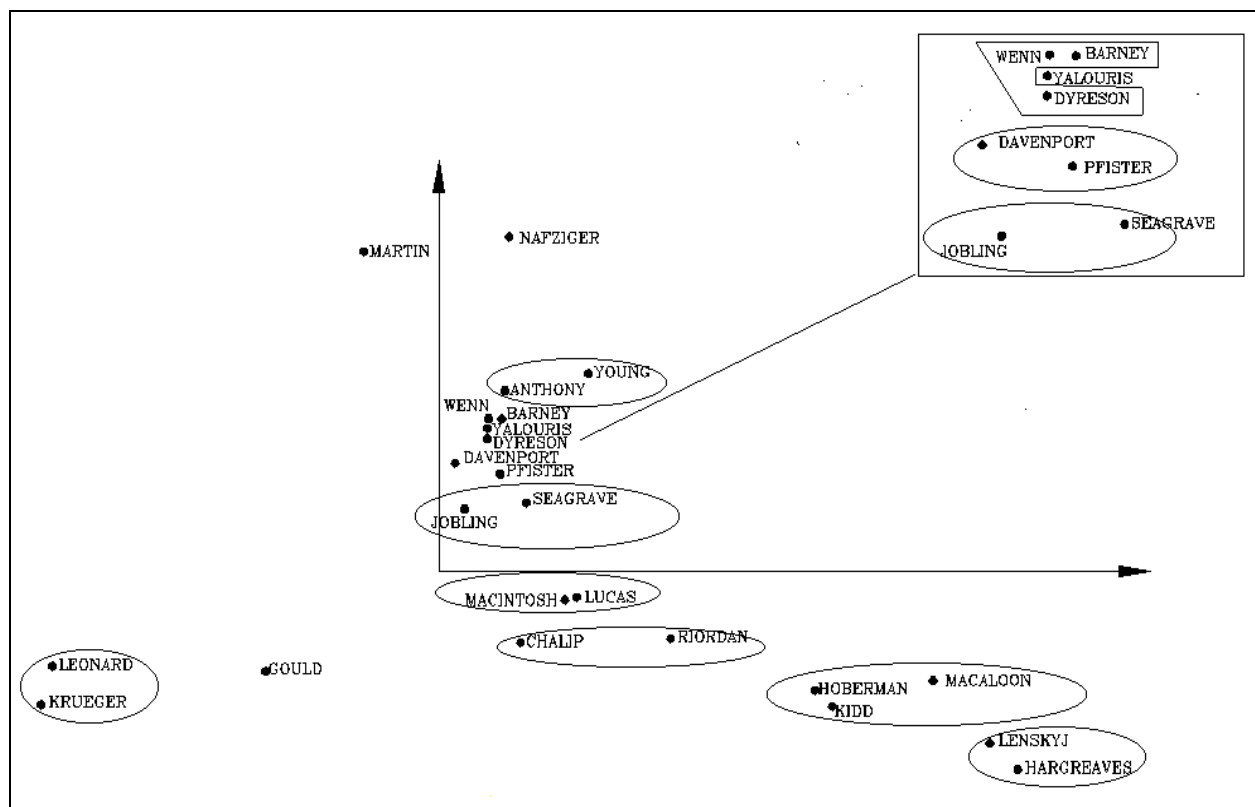
Once the clusters have been identified, they need to be labelled. This was achieved by combining the subject descriptors assigned by indexers in the *Sport* database to the authors' articles with the subject expertise of the field expert.

So for example, for author Riordan table 2 depicts some of subject headings and descriptors applied to his articles indexed in the *Sport* database.

Table 2
Selection of subject headings and descriptors for author Riordan

Olympic Games, Moscow 1980
women
People's Republic of China
doping
sociology
history
sex factor
policy
diplomacy
communism
politics and government - ideology
Eastern Europe
ideology
sociology - social change
Olympism
Olympic Games - philosophy
etc

4. Limitations

Like most techniques, there are some problems and grey areas. For example, as indicated above, false drops may occur when both authors being tested have common names. Co-authorship also has its problems. Only the first named author will appear as a cited author due to the way *Social SciSearch* is organised, so co-authors will miss out. Authors who are cocited rarely with others in the group may skew the statistical results, so they have to be placed manually (on the basis of field knowledge) on the map or excluded altogether.

Selection of the author group may be done in a number of ways. It can be done in a purely objective way, based on gross publishing output or gross citation output. A combination of these two objective methods can be used. The author list can be restricted to certain influential journals. Authors can be included on the basis of perceived importance by a subject expert or experts. Combinations of these approaches are numerous.

There is also the selection of the time frame to consider. It has to be long enough to enable a sufficient body of work to be generated and cited. It should not be too long so that trends or evolution in the literature become clouded. White & McCain (1998) divided the period from 1972-1995 into three 8-year periods in an attempt to observe and measure the way scholarship had changed in the discipline of Information Science. If the period is too recent, it may not allow sufficient time for scholarship to be picked up, digested and then cited in subsequent publications due to the publishing lag.

Settling on the time period for Olympic scholarship has some complications due the quadrennial nature of the Olympic Games, which drives much of what is written about them. The authors settled on a 10-year period, covering two Olympic Games, plus 2 years before and 3 years after the quadrennials. There is nothing magic about the 10-year period, but it does have a certain roundness to it.

The statistical treatment of the cocitation results and their interpretation by the authors provide the following clusters, labelled by the field expert:

Lenskyj and Hargreaves (Critical feminists)
Although separated by the Atlantic Ocean Helen Lenskyj' and Jennifer Hargreaves' work has much closer proximity in their theoretical underpinnings. Both write from a critical feminist viewpoint about gender inequities in the Olympic Movement and proposals for an agenda of reform.

MacAloon, Hoberman and Kidd (Critical reformers)
The authors in this cluster have written on a variety of Olympic related topics, including the Olympic ceremonies, cultural festivals, politics and women. Much of their writing, while critical of the Olympic Movement, also offers suggestions for reform.

Riordan and Chalip (Sport policy and international relations)
Much of the writing of James Riordan and Lawrence Chalip deals with national and international policy in sport in general. Their scholarship on the Olympic Games forms a subset of this.

MacIntosh and Lucas (Ideals and Questions)
Sometimes it is hard to understand why authors are grouped in the same cluster, other than the basic premise of both being cited in the same articles. This is one such pairing. Lucas is a traditionalist and idealist, upholding the ideals of the Olympic Movement. MacIntosh, whose Olympic writings deal primarily with issues of politics, is far more censorious of the IOC.

Leonard and Krueger (Drugs)
Some clusters are formed through authors having a shared theoretical viewpoint, other through topics of research. This is the latter, as the commonality in this grouping is the subject matter of performance enhancing drugs.

Anthony and Young (The revival)
These two authors have written extensively on the early days of the modern Olympic Games and the conditions in Europe in the nineteenth century that led to the establishment of the International Olympic Committee.

Gould (Athletic performance)
Gould stands alone as his research is in the area of psychology, coaching and athletic performance.

Nafziger (Legal aspects)
Another non-aligned writer, Nafziger has investigated the law and the Olympic Games.

Martin (Performance)
The third "free agent" on the map, Martin's work has examined factors that have influenced athletic performance at, and the results of, the recent Olympic Games.

Pfister and Davenport (The history of women's Olympic involvement)
Both writers focus on the history and sociology of the battle for women to achieve equality in the Olympic Games.

Yalouris (The Ancient Games)
It is fitting that Yalouris is located towards the centre of the map, as his own writings are on the topic of the Ancient Games, the cornerstone of the Modern Games, and he has edited, for many years, the proceedings of the International Olympic Academy, the educational arm of the International and Hellenic Olympic Committees.

Wenn, Barney, Dyreson (the North American perspective)
While the topic matter of these authors varies, their linkage has occurred through their writings on North America and that continent's relationship with and effect on the Olympic Games.

Seagreave and Jobling (Olympism)
These authors' citations deal with a variety of Olympic related research. One topic in common is their examination of Olympism, the philosophy of the Olympic Movement.

6. Conclusion

While this investigation has constructed an interesting, and hopefully, insightful, overview of Olympic studies, it does not provide the full picture. Rather, it provides a significant starting point. The research process has yielded clusters of Olympic scholars, based on how others cocite their works.

These clusters have been shaped by different justifications. Some are formed by authors who write on a similar topic (e.g. Chalip and Riordan), others constitute researchers who write from the same theoretical foundation (e.g. Hargreaves and Lenskyj).

Some research areas do not appear prominently as clusters on the map, yet they provide rich and prominent fields of inquiry in sport and physical activity. Two examples are the fields of exercise physiology and sports psychology. There is only one representative from each (Martin and Gould). While no definitive reason can be attributed to this perhaps it is the nature of athletic performance itself that has been a contributing factor. The Olympic Games represent the pinnacle of athletic achievement. Athletes need to be focused on their performance and it is understandably difficult to conduct research on them during Games time. It is far less intrusive to investigate the more macro social and cultural aspects of the Summer Olympic Games during the sixteen days each four years that they are celebrated; the same premise holds true for the Winter Games. Of course, between Games, the same issues do not apply.

It is evident that more research on this topic would be beneficial. A first step would be to enlarge the author list, and refine the methods for selecting the authors, to represent influence, as indicated by numbers of citations rather than gross publication rates. At the other end of the process, some techniques for identifying areas that do not become evident from ACA could be explored.

McCain K. (1990) - *Mapping Authors in Intellectual Space: a Technical Overview.* In Journal of the American Society for Information Science 41 (6), pp. 433-443.

Small H. (1973) - *Co-citation in the Scientific Literature: a New Measure of the Relationship Between Two Documents.* In Journal of the American Society for Information Science 24 (4), pp. 265-9.

Toohey K, Warning P. (1994) - *Mapping the Discipline of Physical Education Pedagogy.* In Access to Active Living: Proceedings of the Xth Commonwealth & International Science Congress, Victoria, Canada, August 1994, pp. 123-7.

Toohey K., Veal A. (2000) - *The Olympic Games : A Social Science Perspective.* Oxford; CABI Publications.

White H., Griffith B. (1981) - *A Co-citation Map of Authors in Judgment and Decision Research.* In B. F. Anderson et al (eds), Concepts in Judgment and Decision Research: Definitions, Sources, Interrelationships, Comments New York; Praeger, pp. 261-271.

White H., Griffith B. (1981a) - *Author Cocitation; a Literature measure of intellectual Structure.* In Journal of the American Society for Information Science 32 (2), pp. 163-171.

White H., McCain K. (1998) - *Visualizing a Discipline: an Author Co-citation Analysis of Information Science, 1972-1995.* In Journal of the American Society for Information Science 49 (4), pp. 327-355.