

Reviewer's report

Title: The intellectual structure and substance of the knowledge utilization field: A longitudinal author co-citation analysis, 1945-2004

Version: 1 **Date:** 4 April 2008

Reviewer: Katherine W. McCai

Reviewer's report:

General Comments:

I should begin by saying that I am a bibliometrician who has done considerable mapping of scholarly literatures – both single period and in time series – but I haven't used the software that the authors employ in their study. I have a basic, not an extensive, knowledge of the work in diffusion of innovation/technology transfer. My comments should be read according.

This is an interesting, well-written article that presents the intellectual structure and research trends in the field of knowledge utilization over a 50 year period—1945-2004. The growth and changes in the field are discussed at length and are well-documented (I leave it to the journal editors to decide whether the article needs cutting – I learned a great deal from the extended discussion.)

The mapped data come from 5278 articles published between 1965 and 2004 that were retrieved from Thomson Scientific's World of Science using a carefully developed, comprehensive search strategy and post-retrieval screening for relevance. The results include both "standard" citation analysis results (lists by decade of most prolific journals, most highly cited authors and works, etc.) and author cocitation maps by decade, using Persson's BibExcel software and Systat. The retrieval method is well-documented—a full search strategy is included in the appendix.

The MDS results could be better documented in terms of giving the readers sufficient information to assess the quality of the maps and compare these results to those examples cited in the literature—particularly since the authors state (MS pg. 9) that they are using ACA "in the manner of White and McCain" but do not follow White & McCain's methodological approaches to data analysis and visualization (see specific comments below).

The results of the study are likely to be of interest both to researchers in the field of knowledge utilization and to students of growth and change in scientific fields more broadly. The emergence of EBM in the later years is intriguing and worth additional exploration.

Major Compusory Revisions

1. The authors are mistaken about the content of the journal Special Libraries (in MS pg 19). Unless there were two journals by the same name, Special Libraries

was a publication of the Special Libraries Association and the articles reflected the broad range of interests of this group of practitioners (and some researchers). This can be confirmed by examining the articles indexed in SSCA 1972-1974. Its appearance here is intriguing.

Minor Essential Revisions

1. The title is a bit misleading, since the maps are based on a shorter time frame than the full set of article counts. The mapping only covers the 1965+ period although the lists span a longer time frame.

There are a number of small issues relating to the MDS mapping that should be addressed to explain the specific methods used and the quality of the results.

2. How were the diagonal cells defined in the cocitation matrices? There are different views on this.

3. Most ACA maps are based on either Pearson correlations (White, McCain and colleagues) or Salton's cosine (Leydesdorff's and possibly Persson's preference) -- these measures tend to reduce the differences in scale of citedness and focus on pattern similarity = subject structure in the case of ACA and Journal CA at least. Authors with high cocitation pattern similarities are placed near each other, as a rule, and those with similarities to many authors placed in the central region of the map. This is interpretable based on subject or similar pattern relationships. The raw cocitation counts are more commonly visualized using a network analytic tool such as Schvaneveldt's KNOT software (PFNets) which emphasize citedness and centrality in the network [1]. What is the effect of using MDS directly on the raw cocitation counts?

4. Non-metric approaches to scaling yield a higher % variance explained with a lower Kruskal Stress I (the measure of goodness of fit reported by the authors) than do metric approaches – which approach was used in the SYSTAT MDS mapping?

5. Reporting the % variance explained by the two-dimensional solution is important as well—a low % variance explained can suggest either the need for a higher dimensional solution or a problem in coding the data. In ACA studies, % variance explained below ~0.8 are generally not considered good except in unusual circumstances. What was the % variance explained in each map?

Discretionary revisions

Methods again

1. The authors identified the author groupings in the maps based in inspection and the “highly cited/productive” lists rather than using cluster analysis or factor analysis (principal components analysis) or some other pattern extraction program independently on the data. This is reasonable but leaves out the opportunity for detecting additional interesting patterns and relationships and introduces a bit of additional subjectivity. The authors may want to comment.

2. No explanation is provided for the need to keep the 1965-74 mapped author set to a count of 13 other than “lack of interpretability” with more (or fewer)

authors. If the problem was noise introduced by large numbers of very small cocitation counts, that's one thing, but if there was structure that needed more exploration, then it would be good to present/discuss this. Reification is always a danger when dealing with these kinds of analyses and unexpected results can illuminate interesting aspects of a field.

Results & Discussion comments

3. With respect to Bradford Zones, while Bradford reported 3 zones, Brooks [2-3] later showed that the same article set could be partitioned into several different numbers of zones. How important is the existence of three zones?

4. When Knowledge changed to Science Communication, it not only changed editorship but the orientation of the content as well. This likely accounts for its post-1994 disappearance from the core list.

5. While White is certainly correct that the ACA maps are to be considered in the "light of the claims being made" it is also the case that these maps have been quantitatively validated as being quite congruent with data collected independently from scholars in the field (as is demonstrated in MS reference #22 and elsewhere).

6. A limitation of the study that is not mentioned is the effect of partitioning on papers/authors/journals published toward the end of the decade. These have a lower likelihood of being cited than do those published toward the beginning of the decade. There's no good way around this without resorting to some kind of continuous interactive display, but it should be mentioned.

1. White, Howard D. (2003). Pathfinder Networks and Author Cocitation Analysis: A Remapping of Paradigmatic Information Scientists. *Journal of the American Society for Information Science and Technology* 54: 423-434.

2. Brooks, TA. Literature Core Zones Adjusted by Impact Factors. *Journal of Information Science*, v. 16 (1), 1990, 51-57.

3. Brooks, TA Clustering in Comprehensive Bibliographies and Related Literatures." *Journal of the American Society for Information Science*, v. 41 (3), April 1990, 183-192.

Level of interest: An article of outstanding merit and interest in its field

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I declare that I have no competing interests.