Information Epidemics and the Transformation of Science

Christine Dufour and Albert N. Tabah
École de bibliothéconomie et des sciences de l'information
Université de Montréal
dufourch@ere.umontreal.ca
tabahan@ere.umontreal.ca

The purpose of the present paper is to characterize the literatures of information epidemics, to analyze how the intellectual structure of a research specialty is transformed as a result of a major discovery, and to describe how scientists respond to rapidly changing conditions in their specialty. A recent study of the physics literature in the 1970s and 1980s has identified high temperature superconductivity research as one such epidemic. A citation and cocitation analysis for the years 1985, 1987, 1989 and 1991 illustrates the structural transformation of the specialty.

Le but du présent travail est de caractériser la littérature des épidémies d'information, d'analyser la transformation d'une spécialité suite à une découverte majeure et de décrire comment les scientifiques réagissent aux changements rapides dans leur domaine. La littérature sur la supraconductivité a récemment été identifiée comme une épidémie d'information. Une analyse de citations et de cocitations pour les années 1985, 1987, 1989 et 1991 illustre la transformation de ce domaine scientifique.

Introduction

In one of the most influential works of the twentieth century, Thomas Kuhn has popularized the idea that progress in science consists of small steps and quiet work (which he termed "normal science") interspersed by occasional scientific revolutions that result in paradigm shifts. From the information science perspective, many of these revolutions are accompanied by a sudden and dramatic growth of their scientific literatures. These sudden growth patterns are termed information epidemics (Tabah 1996). These epidemics often arise subsequent to the publication of an influential work that draws a large number of scientists into a field and provokes them to publish extensively. Once the growth takes place, it can be followed either by a decay process due to the stalling of progress or else it can result in a state of permanence due to continued progress in the field. The continued success reflected in the high rate of publication is accompanied by structural changes in the field that are reflected in the cocitation patterns.

A recent study of the physics literature in the 1970s and 1980s has identified a number of research specialties that have exhibited such information epidemics (Tabah 1996). One of these, high-temperature superconductivity (HTSC), represents the classic case of a scientific revolution. In a surprising paper in late 1986 Johann Bednorz and Kurt Muller claimed that they had achieved superconductivity at 30° K in a compound based on copper oxide and rare earth metals (Bednorz and Muller 1986). The best that had been achieved until that time had been 23.3°K in some metal composites of so-called A-15 compounds. Almost immediately, researchers put aside their experiments, verified the veracity of the information and switched to the family of copper oxide-based compounds. The flurry of activity during the six months subsequent to the publication of the paper raised the critical temperature past the 90°K mark and transformed the field. Large numbers of researchers joined superconductivity research programs, and a number of Western countries contributed new funds to high-temperature superconductivity research (OTA 1990).

Our approach so far has been to identify and characterize the literatures of information epidemics, to analyze how the intellectual structure of a research specialty is transformed as a result of a major discovery, and to discover how scientists respond to rapidly changing conditions in their specialty. While physics, and particularly superconductivity, have been popular subjects for scientometric studies in the past, most published works have been restricted to work conducted during the 1960s and 1970s. As to work on HTSC, Brooks in 1990 conducted an analysis of the journals publishing HTSC articles, and Nowotny and Felt gave an overview of the impact of the discovery on the science system (Brooks 1990; Nowotny and Felt 1992). However, no one has studied the intellectual structure of superconductivity following the changes immediately after the publication of the Bednorz and Muller paper. The purpose of this paper is to further report on recent work detailing the intellectual and the structural changes that took place immediately before and after the information epidemic.

Methodology

The Analysis Tool

The analysis was conducted with *BiblioMétrika,* a tool recently developed for our own analyses, that makes full use of the capabilities of *Access*, a relational database available from *Microsoft Inc.* It has been designed to take into consideration the characteristics of *INSPEC* and *Science Citation Index* (SCI) records, the work flow, and the bibliometric analyses to be performed.

Figure 1. Workflow

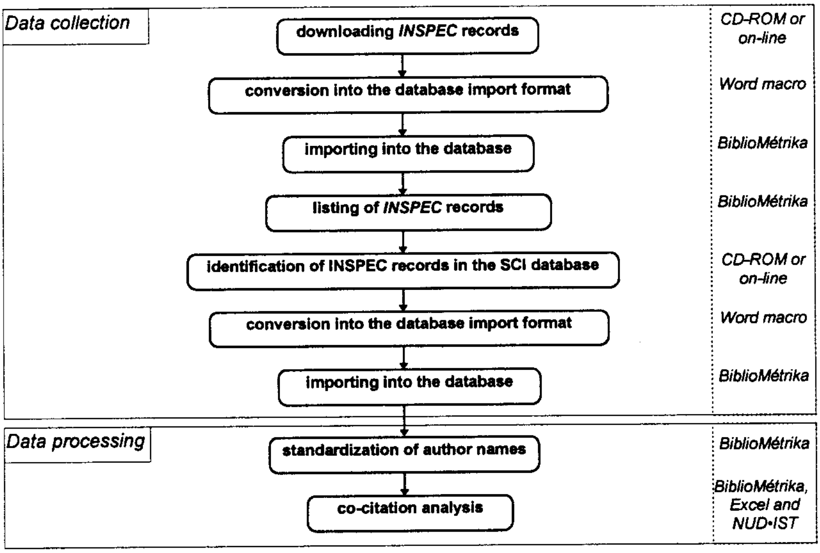


Figure 1 illustrates the data processing and the data collection steps undertaken for this work. The right-hand column of the figure indicates the tools used during the implementation of each step.

The initial data collection phase involved several steps. The downloaded sets of *INSPEC* records were converted through a *Word* macro into the database format. Listings of source records were printed for use in the identification of corresponding SCI records. Once identified and downloaded, SCI records were converted into the *BiblioMétrika* format.

The data processing phase, at first, consisted of standardizing the author names so as to increase the data quality. This implies human judgment and cannot be automated; *BiblioMétrika* can only facilitate the manual work by supplying adequate interfaces. This standardization mechanism results in a controlled vocabulary indicating accepted and rejected forms. The current version of *BiblioMétrika* can be used to generate frequencies (authors, journals, citations, descriptors or classification codes) and to conduct further analyses such as cocitation or coword analyses with other computer packages. For example, tables or matrices can be fed into *SPSS* or *Statistica* to generate clusters or maps.

Data Analysis

A multidimensional scale analysis of articles published in *Physics Abstracts* and subsequently identified in the *Science Citation Index* was performed. All the articles covered in Chapter 74 (Superconductivity) of *Physics Abstracts* for 1985, 1987, 1989 and 1991 were identified in SCI and downloaded. A citation and cocitation analysis were performed on each year of the data. The choice for these years was motivated by the need to illustrate the state of superconductivity research preceding and following the publication of the Bednorz and Muller article. The use of two-year intervals adequately reflects the evolution of superconductivity research. We have previously reported on work between 1985 and 1989 (Tabah and Dufour 1997). This contribution extends the analysis to 1991 and refines the methodology.

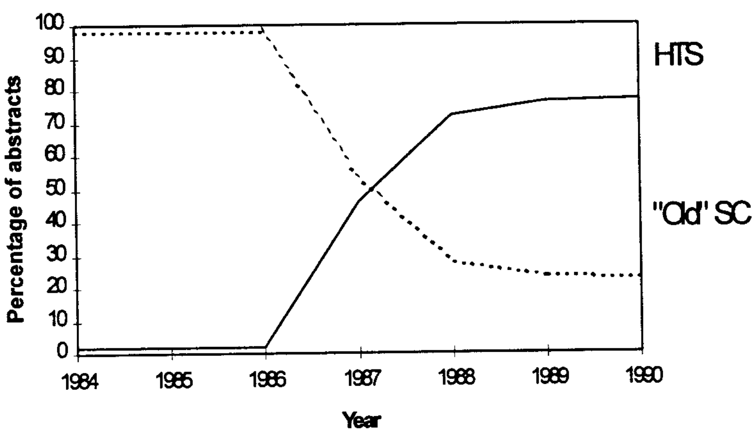
The data analysis followed the methodology developed by White and Griffith and by McCain (White and Griffith 1981; McCain 1990). After eliminating self-citations, authors were ranked by citations received and all those above the 0.30% threshold were searched for cocitation relationships. The cocitation matrix was fed into *Statistica* and converted into a correlation matrix. Author groups were formed using Ward's method with complete linkages in the "Cluster Analysis" module of *Statistica* and converted into MDS maps using the

"Multidimensional Scaling" module. The cluster cut-offs were determined with the help of the "Amalgamation Schedule", akin to a scree test in factor analysis. The first pages of the most cited articles forming each cluster were examined, significant words were entered into NUD•IST (a qualitative analysis tool) and the cluster names were derived from the frequency matrix.

Results

The phenomenal growth of publications following the Bednorz and Muller article has been reported previously (Tabah, 1995). By late 1987, the number of papers indexed by Chapter 74 (Superconductivity) of Physics had jumped fourfold from an average of 75 per month in 1985-1986 to about 300 in 1988-1989. Almost all of the new publications were on high-temperature superconductivity (HTSC) research. By the beginning of 1988, that is within a little over a year, 72.4% of superconductivity work reported in the literature had been carried out on the new HTSC related materials (see Fig.2 below). The remaining publications either belonged to work continued on previously used materials or else were indexed late by *Physics Abstracts* due to certain time lags inherent in publishing and indexing. In order to differentiate the new work conducted on the family of cuprate oxide materials from work conducted prior to the publication of the Bednorz and Muller article, the former have been termed HTSC work and the latter "Old" SC work.

Figure 2. Transition from the Old to the New Superconductivity.



Figures 3-6 below show the mappings generated for the years 1985, 1987, 1989 and 1991. Although the number of authors jumped considerably from 1985 to 1991, the number of journals publishing their articles remained relatively constant. In our case, the number of authors in 1985 was 1328, in 1987 3152, in 1989 5055, and in 1991 5214. This represents an almost fourfold increase from 1985 to 1989. The number of journals has been covered previously by Brooks and need not be repeated here (Brooks, 1990).

Figure 3. Cocitation Map of 1985.

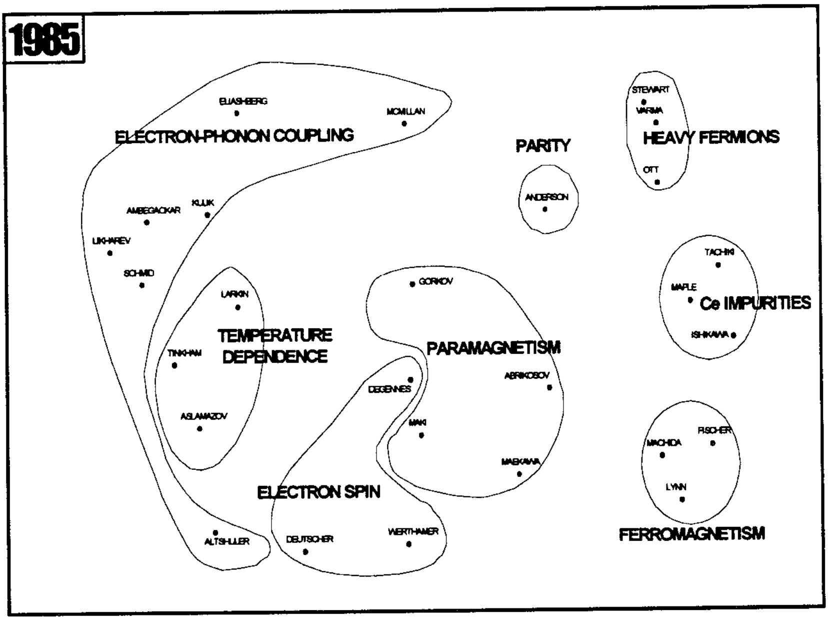


Figure 4. Cocitation Map of 1987.

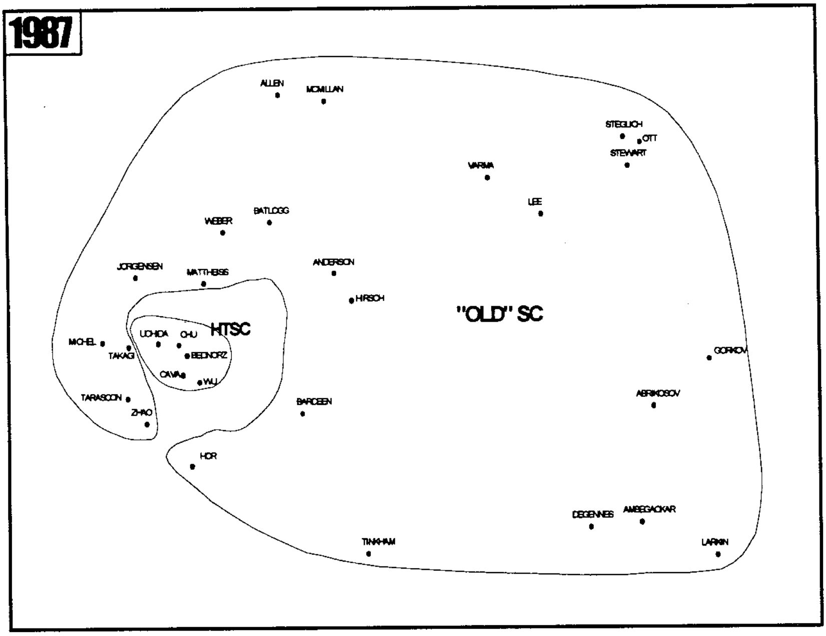


Figure 5. Cocitation Map of 1989.

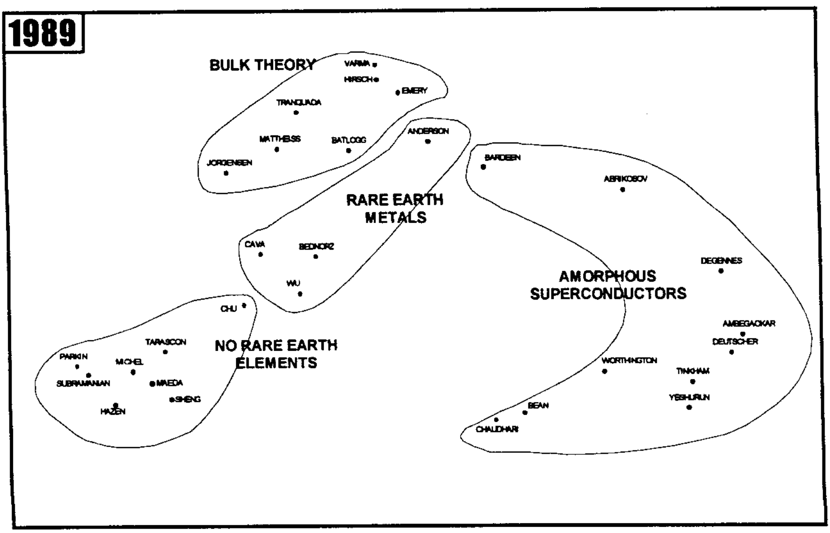
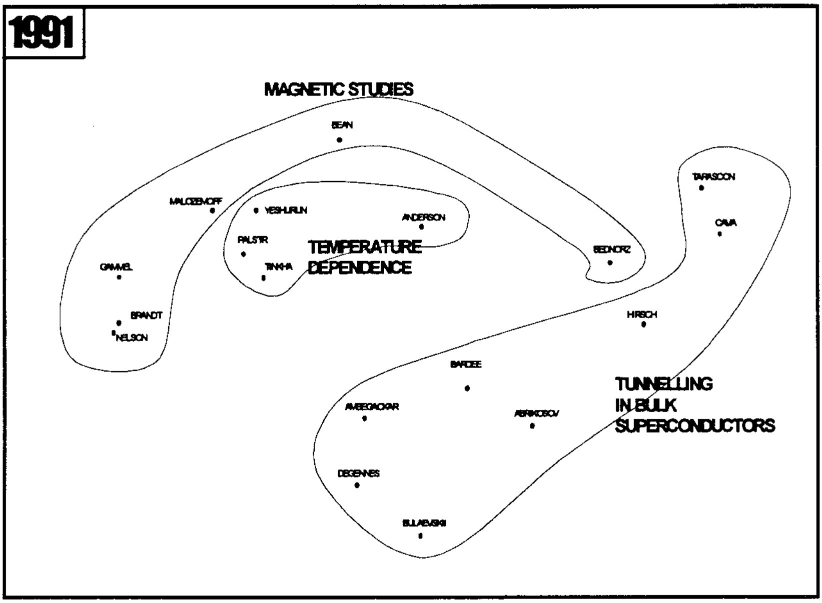


Figure 6. Cocitation Map of 1991.



Discussion

The four maps presented so far adequately reflect the transformation of superconductivity research following the epidemic. From the 1985 map it is difficult to define a common approach. The horizontal axis of the map (the axes are not shown for the sake of clarity) seems to be based on the extent of impurities in the test compounds whereas the vertical axis differentiates them by microscopic effects. The left-hand side of the map is more concentrated on impurity-free compounds, whereas the right-hand side is working on the introduction of impurities into bulk materials. The lower part of the vertical axis is dealing with electron-spin whereas the upper part is dealing with tunneling effects. This is consistent with the state of superconductivity research in the mid-1980s (Allen and Mitrovic 1982).

In 1987 a clear delimitation begins to emerge: a small island of HTSC workers introduces itself into a large sea of authors working on older materials. There is probably some artificiality introduced here by publishing and indexing lags. The average number of citations received in 1987 by the authors forming the HTSC cluster is 274 whereas their average cocitation strength is 149 (ten times the average of 14.3). The lower part of the vertical axis represents those working on temperature dependence whereas the upper side represents those working on heavy fermion elements such as platinum and uranium.

The horizontal line formed by Ambegaokar-Larkin-Tinkham-DeGennes in the middle of the 1985 map has been pushed south and has formed a new base. The McMillan-Stewart-Ott-Varma line in the upper side has retained its position. In-between, from the left-hand side, has come the new HTSC group composed of Bednorz, Cava, Chu and Wu.

In 1989, the gap has increased considerably. The HTSC group of workers has invaded the majority of the space on the map, and the "old" SC workers have been pushed to the periphery. Of the four clusters, three deal with HTSC materials. Bednorz, Cava, Chu and Wu have taken the center stage. The only differentiation among them now is that one group on the lower left, including Chu, works on materials with no rare earth elements whereas another, at the center, led by Bednorz, Cava, and Wu still works on materials that include rare earth elements. The group characterized by "no rare earth elements" now represents a maverick group in itself, in that several of its members have gone beyond the initial family of materials discovered by Bednorz and Muller. In fact, in 1988 two of its members discovered superconductivity above 100°K, one being Maeda with Bi-Ca-Sr-Cu-O and the other being Sheng with Tl-Ca-Ba-Cu-O (Maeda et al. 1988; Sheng et al. 1988). A third cluster, at the top of the map, works on bulk materials and is composed of authors who have also switched their attention to new HTSC materials. Finally, the fourth on the right-hand side is composed of authors representing the interests of research on older type materials.

In 1991 the superconductivity map consists exclusively of HTSC workers. What began as a surprising discovery not only gained an immediate foothold but completely overtook superconductivity research within four years. The "old" SC group of 1989 has disappeared, but some of its members have survived in the tunnelling group. These, such as Abrikosov and Ambegaokar, are researchers whose basic work years ago still serves to put recent work into perspective. Three major names occupy central positions: Bardeen due to his BSC theory, Anderson due to his theoretical prominence, and Bednorz due to his ground-breaking work. On the other hand, the common thread among the three clusters identified is the search for theoretical unity. The old BCS theory not being sufficient to explain modern HTSC phenomena, the need arises to find a new theoretical base. One such possibility is offered by the tunnelling group that studies the properties of two-dimensional layers in superconductive crystals.

In most comparative cocitation studies encountered in the literature so far most individuals within groupings have remained stable from one comparison year to another. In our case the author groupings are rather unstable; from one period to another very few authors have kept their positions. The notable individuals are Anderson and Bardeen, both centrally located, confirming their theoretical hold on the field. In addition, certain author groupings managed to survive: Stewart-Ott-Varma (from 1985 to 1987), Jorgensen-Mattheiss-Batlogg (from 1987 to 1989), and Bednorz-Cava-Chu-Wu (from 1987 to 1989). On the other hand, in 1991 almost all of these groupings have disappeared. Of the original HTSC pioneers only Bednorz and Cava remain. Very few authors are present in all the maps presented: Abrikosov, Ambegaokar, Anderson, DeGennes, and Tinkham. This is most likely due to the fact that they all represent classical studies on which superconductivity work is based.

All of these authors are prolific, and an analysis of their output shows that their publications cited during this time form coherent clusters, clearly differentiating them from one another. The 1985 map is characterized by eight clusters, and it is very difficult to see a uniform approach toward research in superconductivity in this period. Subsequent to the publication of the Bednorz and Muller article in 1986 the number of clusters in the 1987 map is reduced to two, just at the time when HTSC makes its entrance into the field. By 1989, HTSC is well-established, it has an entrenched position, and the map is divided into four clusters. In 1991, the takeover by HTSC is complete and the clusters are reduced to three. Thus, over the six years and four maps of analysis, superconductivity undergoes a transformation that is reflected in the number of clusters themselves: a lack of coherence in 1985, a sudden fusion that distinguishes the old from the new in 1987, a re-division of interests in 1989, and a stabilization in 1991. This also conforms to the model of scientific growth of Lemaine (Lemaine et al. 1976).

The time spans covered by the cited publications differ among the clusters over the years analyzed. The time span of works for cited authors within the 1985 clusters covers 32 years. There are several works going back to the 1950s and early 1960s. In 1987 there is a wide difference between the two clusters. The publications of HTSC authors all originate from a four-month span in early 1987 whereas those of "Old" SC authors go back to 1950. By 1989, the time span covered, especially for the three groups working on HTSC is barely two years whereas that for the older group represented by the "amorphous superconductors" cluster is up to 37 years. In 1991, the time span for the groups still goes back to 1946, in reference to the theoretical basis of superconductivity. Progressing from 1985 to 1991, the median age for citations is halved during the exponential phase of the epidemic: 8 years for 1985, 4 years for 1987 and 2 years for 1989, and then extends to 5 years once the growth subsides. This reflects the rapidity with which scientists worked and published during the flurry of the initial excitement and typifies what is expected of an information epidemic. The table below summarizes the relevant statistics for the three maps:

Table 1. Characteristic Statistics for the Three Years of Analysis.

	1985	1987	1989	1991
Number of authors	27	30	29	18
Stress	11.4%	14.3%	7.5%	9%
Median citation age (range of years)	8 (1953-85)	4 (1950-87)	2 (1950-89)	5 (1946-90)
Mean cocitation rate (range)	4.5 (0-25)	14.3 (0-220)	15.5 (0-175)	29.2 (0-72)
Number of clusters	8	2	4	3

Conclusions

Overall, the behaviour of the rapidly growing field of high-temperature superconductivity is entirely consistent with the characteristics of fast-growing literatures and information epidemics: rapid impact, high citation frequency, swift diffusion, a small group of fundamental papers attracting new workers into the field and the subsequent swelling of author numbers (Tabah 1996). At the same time, it reflects the fierce competitive efforts among scientists to be one of the first to discover a compound at a yet higher critical temperature and perhaps even a new compound that will constitute the basis for a new family of materials. Work continues on extending the analysis and adding new case studies to the thesis of information epidemics.

References

Allen, P.B., and B. Mitrovic. 1982. "Theory of superconducting Tc," *Solid State Physics* 37: 2-92.

Bednorz, J.G., and K.A. Muller. 1986. "Possible high Tc superconductivity in the Ba-La-Cu-O System," *Zeitschrift fur Physik* 64: 189-193.

Brooks, T.A. 1990. "Core journals of the rapidly changing research front of superconductivity," *Scholarly Communication and Bibliometrics*. ed. C.L. Borgman, 235-247. Newbury Park, CA: Sage Publications.

Felt, U., and H. Nowotny. 1992. "Striking gold in the 1990's: The discovery of high-temperature superconductivity and its impact on the science system," *Science, Technology & Human Values* 17: 506-531.

Lemaine, G. et al., eds. 1976. *Perspectives on the Emergence of Scientific Disciplines*. Paris: Mouton.

Maeda, H. et al. 1988. "A new high-Tc oxide superconductor without a rare earth element," *Japanese Journal of Applied Physics*, Part 2 27: L209-211.

McCain, K.W. 1990. "Mapping authors in intellectual space: A technical overview," *Journal of the American Society for Information Science* 41: 433-443.

Sheng, Z.Z., and A.M. Hermann. 1988. "Bulk Superconductivity at 120°K in the Tl-Ca/Ba-Cu-O System," *Nature* 332: 138-143.

Tabah, A.N., and Christine Dufour. 1997. "The Invasion of the oxides: Transformation of superconductivity research as a result of an epidemic," In *Proceedings of the 6th Conference of the ISSI, Hebrew University of Jérusalem, 16-19 juin 1997*, pp. 439-448.

Tabah, A.N. 1995. "Knowledge epidemics in the growth of physics," In *Proceedings of the Fifth Biennial Conference of the International Society for Scientometrics and Informetrics, Chicago, June 7-10, 1995*, edited by M.E.D. Koenig and A. Bookstein, 555-564.

Tabah, A.N. 1996. Information epidemics and the growth of physics. Diss., McGill University.

U.S. Office of Technology Assessment. 1990. *High-temperature superconductivity in perspective*. Washington, DC: OTA.

White, H.D., and B.C. Griffith. 1981. "Author cocitation: A literature measure of intellectual structure," *Journal of the American Society for Information Science* 32: 163-171.