Toward a Consensus Map of Science

Richard Klavans
SciTech Strategies, Inc., Berwyn, PA 19312. E-mail: rklavans@mapofscience.com

Kevin W. Boyack
SciTech Strategies, Inc., Albuquerque, NM 87122. E-mail: kboyack@mapofscience.com

A consensus map of science is generated from an analysis of 20 existing maps of science. These 20 maps occur in three basic forms: hierarchical, centric, and noncentric (or circular). The consensus map, generated from consensus edges that occur in at least half of the input maps, emerges in a circular form. The ordering of areas is as follows: mathematics is (arbitrarily) placed at the top of the circle, and is followed clockwise by physics, physical chemistry, engineering, chemistry, earth sciences, biology, biochemistry, infectious diseases, medicine, health services, brain research, psychology, humanities, social sciences, and computer science. The link between computer science and mathematics completes the circle. If the lowest weighted edges are pruned from this consensus circular map, a hierarchical map stretching from mathematics to social sciences results. The circular map of science is found to have a high level of correspondence with the 20 existing maps, and has a variety of advantages over hierarchical and centric forms. A one-dimensional Riemannian version of the consensus map is also proposed.

Introduction

There has been a great deal of interest in visualizing the structure of science over the past five years. In 2003, the U.S. National Academy of Sciences convened a conference specifically on mapping science (Shiffrin & Börner, 2004). Katy Börner, one of the conference organizers, followed up this effort with a traveling exhibit of science maps (called Places & Spaces) that has appeared at over 50 international locations since 2005. This exhibit (an online version is available at http://www.scimaps.org/) reflects the work being done by research groups from around the world, representing a variety of academic disciplines and using a variety of techniques and databases. Maps from this exhibit have found their way into the permanent map collection at the New York Public Library and in the year-end edition of *Nature* (Marris, 2006).

Given the number of science maps that have appeared in the literature with increasing frequency, we wondered whether these maps are starting to converge on a common solution, or if a consensus among maps was being formed. We differentiate here between consensus and convergence; they are two very different things. If convergence is occurring, all recent maps that look at a similar slice of science (e.g., all of science) should look nearly the same in terms of form, content, and linkages. Consensus is a lower standard, and implies that an aggregation of results from a variety of input maps would share a large number of common features with the individual maps.

A review of the literature has shown that convergence in science maps is not occurring. However, we felt that consensus was very possible. A consensus map, if it exists, would be extremely helpful in the adoption and application of science maps. A consensus map could be useful as a teaching aid in elementary and secondary education. A consensus map can raise the general awareness of the importance of science and provide a common cognitive framework for the discussion of science policy issues, such as fundamental changes in interdisciplinary relationships. More importantly for this work, a consensus map can help highlight fundamental differences in the complex maps that are being proposed by researchers, and suggest how those differences might be bridged.

In this paper we examine and codify 20 existing maps of science in an attempt to see if there is a consensus that is forming. The paper proceeds as follows. We first set the stage by discussing differences between classification, science mapping, and knowledge mapping, and address the specific criteria used to qualify an existing work as a map that could be used as input for this study. A brief description of each of the 20 maps of science that were selected for inclusion in the study is then given. We then generate a list of high-level disciplines that seem to be common to the majority of the 20 existing maps, and make them (and possible linkages

Received January 28, 2008; revised September 22, 2008; accepted September 22, 2008

 • Published online 8 December 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asi.20991

between them) the basis for looking for consensus. Although somewhat subjective, this is a necessary step to searching for consensus, due to the fact that maps are generated from a variety of different data sources at different levels using different techniques. Each of the 20 maps is then codified in terms of this basis set of high-level disciplines; each map is reduced to a set of disciplines and the relationships between them. A consensus map is then generated from these data, and the consensus map is compared to each of the 20 input maps in a quantitative manner.

The consensus map that emerges from these data is circular (or noncentric) in form. We conclude the paper with a discussion of the reasons for adopting a noncentric model of the structure of science, and a summary of our findings.

Classification, Science Mapping, and Knowledge Mapping

Before considering the existing models or maps that form the input for this study, it is useful to take a step back and define what a science map is and what it is not. This requires a discussion of classification and the differentiation of a science map from a knowledge map.

Classification of science into partitions dates back into the early 19th century, at least to the time of August Comte. Comte not only named six fundamental sciences (i.e., classification), but also placed them in an ordered hierarchy (i.e., a map):

> As a definitive result, mathematics, astronomy, physics, chemistry, physiology, and social physics; such is the encyclopedic formula which, among the great number of classifications which the six fundamental sciences include, is solely in logical conformity with the natural and invariable hierarchy of phenomena. (Comte, 1830, p. 115)

That such efforts have always had critics was as true in the 19th century as it is today. For example, Herbert Spencer, an English philosopher and social theorist, while not disagreeing with the fundamental sciences named by Comte, was highly opposed to Comte's mapping of those sciences:

> From our present point of view, then, it becomes obvious that the conception of a serial arrangement of the sciences is a vicious one. It is not simply that the schemes we have examined are untenable; but it is that the sciences cannot be rightly placed in any linear order whatever . . . There is no one rational order among a host of possible systems. (Spencer, 1864, p. 144)

In general, a map of science consists of a set of elements along with the relationships between the elements. These elements can be scientific fields or disciplines, journals, papers, or any other unit that represents a partition of science. The characteristics that differentiate a map from a simple classification system are (a) the visualization of the elements, commonly represented by locating each of the elements in two-dimensional space, and (b) the explicit linking of pairs of elements by virtue of the relationships between them. From the mapping perspective, classification is often thought of as a step along the way to creating a visual map, but is not equivalent with mapping if the relationships between the classes are not explicitly specified. Maps of science are commonly visualized as node-edge diagrams, similar to those used in network science.

Classification of science, or separation of science into different partitions, is commonly accepted today, and is extremely useful for the cataloging and retrieval of source materials. Among such systems, the U.S. Library of Congress (LOC) has perhaps the most well-accepted classification system in use today (http://www.loc.gov/catdir/cpso/lcco/). However, to the best of our knowledge, the LOC system has not been mapped, meaning that it has not been placed in a visual format where the links between the various category codes are explicitly shown. There are some who would argue that there is an inherent hierarchy to the LOC system that could be mapped as a tree-like structure. For example, class Q (Science) has eleven subclasses (QA-QR, representing disciplines such as mathematics, astronomy, physics, and so forth, and one can assume that each of the subclasses links to the parent class. The difficulty arises in that what we call "all of science" is comprised of at least a half-dozen classes (e.g., medicine, agriculture, social sciences, technology, etc., are separate classes), and there is no explicit linking between classes. This does not reduce the usefulness of the LOC system as a gold standard for classification, but merely means that it does not qualify as a map, using the definition given here. A similar argument can be made for other wonderful resources such as encyclopedias; Britannica's Propaedia while it gives an outline or classification of knowledge, is not a map.

Science mapping, as practiced today, has a far less storied history than classification, and has its roots in the realization that multidimensional spaces can be projected down to two dimensions using multidimensional scaling and related techniques. A multitude of different two-dimensional projections can be derived from the same data due to the use of different similarity measures, algorithms, and projection choices. Consequently, arguments similar to that expressed by Spencer are not uncommon today. When the additional variance from the use of different data sources is added to this mix, the thought of a consensus map of science becomes even more compelling. If a consensus map does exist, overcoming the differences in data sources and mapping variables, it would be a strong indicator of robustness in the high-level structure of science.

In contrast to science mapping, knowledge mapping relies far more on the question of ontology, or what knowledge is and how it might be classified. In addition, knowledge mapping uses a different definition of the word *mapping*. In knowledge mapping, the concept of mapping deals with the correspondence between a classification system and the phenomena in question. In science mapping, the concept of a map draws from cartography. Science maps are analogous to the (hypothetical) floor plan of a library, where books are placed in rooms (i.e., the classification system) and rooms are located so that scholars minimize the distance they have to travel

DOI: 10.1002/asi

(i.e., related areas are proximate). Knowledge maps are sensitive to levels; for example, *infectious disease* can be considered a subset of *medicine*. By contrast, science maps may consider infectious disease and medicine as two categories simply because, as a practical matter in a library, one might have one room devoted to infectious disease and another room with books and journals on other subsets of medicine.

Our application in this paper is entirely related to science mapping. In essence, we are performing a meta-analysis to determine if 20 different maps of the library have common groupings of rooms. Common groupings of rooms in a large majority of maps would indicate that there is a growing consensus in a high-level structure of science. By contrast, this effort has very little to do with knowledge mapping. We are far more interested in the relative placement of the rooms and a summary description of the contents of the rooms. We are not addressing whether the descriptions of these rooms correspond to an ontology of knowledge.

Selection Criteria

Now that we have discussed the differences between classification and mapping, let us set the criteria for including an existing map of science in this study. First, and foremost, it must be a map. Maps conform to the following two criteria: (a) there must be partitions, where science is separated into different parts, and (b) there must be information that links partitions, either through explicit linkages (such as a line drawn between two partitions), or through a combination of proximate location (or physical adjacency on a one- or two-dimensional projection) and accompanying explanation that explicitly states that proximate location denotes linkage. Of course, some maps will have both physical proximity and additional linkages linking areas that are not physically proximate. As mentioned above, neither the Library of Congress classification system, nor the Britannica Propaedia qualify as maps as defined in the previous section.

Second, we focus only on maps that we consider to be comprehensive, meaning that they cover all of science—that is, the physical, biological (including medical), and social sciences—or at least the majority of that space. We note that this is a subjective type of judgment; some maps that we consider to be comprehensive enough for inclusion in the study might be judged otherwise by others.

Basic Map Forms

Before discussing each of the selected existing maps individually, it will be helpful to comment on one of the high-level observations from our codification and analysis of 20 existing maps. Although there are large differences in the complexity of the 20 maps, when reduced to a common level (16 high-level disciplines) we find that there are three basic map forms that emerge (see Figure 1). First, there is a hierarchical form (designated by some authors as a linear model), in which the majority of the disciplines link in a linear sequence. Although there can be a low level of branching in the hierarchical form, the majority of the disciplines are connected by a linear structure. Second, there is a centric form; in this form one discipline lies at the center of a hub-and-spokes type of network in which there is a high degree of branching from the central node. Not all maps with a centric form have the same discipline at the center. The third form is neither hierarchical nor centric, but typically occurs in a ring-like or circular structure. We call this a noncentric form. It differs from the hierarchical form only in that the two ends of the hierarchy are explicitly linked, thus forming a ring. As will be shown for particular cases below, some maps exhibit characteristics of more than one map form.

Selected Maps of Science

Generating a map of science that is relatively comprehensive from bibliographic data is an involved process. One needs a large and highly representative set of data. The data must be highly structured so that parts (clusters of papers, clusters of journals, or disciplines) and the relationship between parts can be adequately modeled. The matrices that

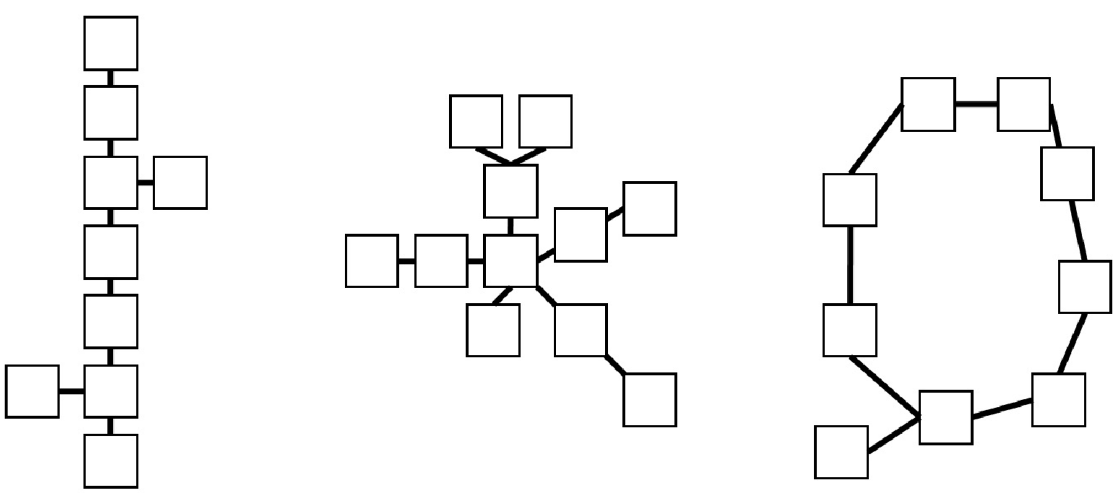

FIG. 1. Examples of hierarchical (or linear), centric, and noncentric (or circular) map forms, from left to right, respectively.

DOI: 10.1002/asi

TABLE 1. Characteristics of 20 comprehensive maps of science. Abbreviations SC, SS, AH, and PR refer to Thomson Scientific's Science, Social Science, Arts & Humanities, and Proceedings Citation databases, respectively.

Researcher(s) & reference	Map name	Method	Elements	# Clust	Database & year	Form
(Bernal, 1939)	Bernal	Expert		14, 110		Hierarchical
(Ellingham, 1948)	Ellingham	Expert		13, 51, 130		Hierarchical & Non-centric
(Balaban & Klein, 2006)	Balaban-I	Expert	16 fields	16		Hierarchical & Centric
(Griffith, Small, Stonehill, & Dey, 1974)	Small74	Reference papers	1,150 papers	41	SC, 1972 Q1	Centric
(Small & Garfield, 1985)	Small85	Reference papers	~ 11,000 papers	51	SC + SS, 1983	Hierarchical & Centric
(Small, 1999)	Small99	Reference papers	36,720 papers	35	SC + SS, 1995	Hierarchical
(Klavans & Boyack, 2008)[a]	KB-Para	Reference papers	800 k papers	776	SC + SS, 2003	Non-centric
(Klavans & Boyack, 2007)	KB06-TS	Reference papers	1.9 M papers	283	SC + SS, 2004	Non-centric
(Klavans & Boyack, 2007)	KB06-SC	Reference papers	2.1 M papers	554	Scopus, 2004	Non-centric
(Bassecoulard & Zitt, 1999)	B-Z	Journals	~ 2,000 jnl	29	SC/JCR, 1993	Hierarchical & Centric
(Klavans, 2002)	K02	Journals	5,647 jnl	69	SC + SS + AH, 2000	Non-centric
(Boyack et al., 2005)	Backbone	Journals	7,121 jnl	205	SC + SS, 2000	Non-centric
(Boyack et al., 2009)	BBK02-S	Journals	7,227 jnl	671	SC + SS, 2002	Non-centric
(Boyack, 2009)	B03-ST	Journals	8,667 jnl	852	SC + SS + PR, 2003	Non-centric
(Klavans et al., 2008)[b]	UCSD	Journals	16,235 jnl	554	SC/SS/AH + Scopus, 2001-05	Non-centric
(Rosvall & Bergstrom, 2008)[c]	Rosvall	Journals	6,116 jnl	87	SC + SS, 2004	Non-centric
(Moya-Anegón et al., 2004)	Scimago-I	Journal categories	25 categ	25	SC + SS + AH, 2000 Spanish papers	Non-centric
(Moya-Anegón et al., 2007)[d]	Scimago-II	Journal categories	219 categ	219	SC + SS + AH, 2002	Centric
(Leydesdorff & Rafols, 2008)[e]	L-R	Journal categories	6,164 jnl; 172 categ	172	SC, 2006	Mixed
(Balaban & Klein, 2006)	Balaban-II	Course prerequisites		11	Texas A&M undergraduate	Centric

[a] http://commons.wikimedia.org/wiki/Image:Topic_map_of_science.jpg
[b] http://scimaps.org/dev/big_thumb.php?map_id=164
[c] http://www.eigenfactor.org/map/maps.htm
[d] http://www.scimago.es/benjamin/USA-2002.jpg
[e] http://users.fmg.uva.nl/lleydesdorff/map06/index.htm

are required can be extremely large (ranging from a few hundred to a few million rows and columns). Methodological compromises are often necessary due to the lack of algorithms that can handle this level of complexity (Boyack, Börner, & Klavans, 2009). There is very little literature showing how these methodological choices and compromises affect the resultant maps.

Due to the time, costs, and difficulties involved, there are relatively few maps of this scope that have been generated. Twenty such maps are listed in Table 1. Each map meets the criteria listed above: Each is a map of science with both partitions and links, and each is comprehensive, covering all or most of the physical sciences, biological sciences, and social sciences.

We have organized the maps in Table 1 by the method used to identify partitions in science. The earliest overall method, expert judgment (with 3 maps), is followed by the earliest computational method, clustering of reference papers. References papers are used as a basis for identifying partitions in science in 6 of the 20 maps. Clustering of journals, where journal clusters are the partitions, was the next methodology used to map science, and accounts for another 7 maps of science. Disciplinary categories, using the Thomson Scientific (TS) journal categories, account for another 3 maps of science. The final map is based on an analysis of undergraduate course prerequisites at an agricultural college in the United States.

If the maps were placed in order based on the date they were generated, one would see a shift from individual to collaborative activity. Before 2000, three of the six maps were generated by individual efforts (Bernal, 1939; Ellingham, 1948; Small, 1999), and two by a pair of researchers (Bassecoulard & Zitt, 1999; Small & Garfield, 1985). Maps generated after 2000 are mostly by research groups. Eight maps are by a group of three researchers in the United States presenting separately (Klavans, 2002; Boyack, 2009), in pairs (Klavans & Boyack, 2007, 2008; Klavans, Boyack, & Patek, 2008), or all three (Boyack, Klavans, & Börner) together (Boyack, Klavans, & Börner, 2005; Boyack, Börner, & Klavans, in press). Two maps are by individuals in a large research group in Spain (Moya-Anegón et al., 2007, 2004), two maps are in one paper by Balaban and Klein (2006), and single maps were generated by two other research groups (Leydesdorff & Rafols, 2007; Rosvall & Bergstrom, 2008).

Following is a summary of the major aspects of each map. Our focus in this review will be on the characteristic shape of each map, which is exemplified by its classification into one of the three forms mentioned above: hierarchical, centric, or noncentric. Designation of a map as hierarchical, centric, or noncentric is based on a combination of comments by the original authors and our interpretation of the actual maps presented in the referenced papers. Although clustering and visualization algorithms will be discussed in some cases to make certain points, we will not provide an in-depth review or comparison of all of the clustering and visualization algorithms used to generate the 20 maps; this information is available in the original literature.

Maps by Experts

We start with two relatively old hand-drawn maps that were comprehensive with respect to the relevant science of their time. Although science today has a different distribution—in the 1940s the physical sciences dominated biology and medicine, today the reverse is true—these older maps are very detailed and well thought out, and deserve to be mentioned. In addition, we find that these older maps have more in common with current science than we would have expected, and we include them to highlight those similarities.

Bernal (1939), uses a 3×2 table-like layout to locate areas of science. The columns correspond to physical, biological, and sociological sectors of science, while the rows correspond to fundamental and technical approaches. Each of the 3×2 regions contains a hierarchical structure of disciplines, and links are drawn between disciplines and labels on the map. The 3×2 layout is not equally spaced. All sectors of science are well represented but one; mathematics did not appear on this map. The physical sciences sector (column) represents almost 50% of the map, and the technical areas (the bottom row, which includes topics such as engineering and the social sciences) also accounts for more than 50% of space on the map.

We considered this map *hierarchical* along two dimensions. Along the x axis of his graph, Bernal clearly shows the dominance[1] of the physical sciences over the biological sciences, and then the dominance of the biological sciences over the social sciences. Along the y axis, he makes the hierarchical distinction between fundamental and applied science.

[1]Here, *dominance* is not meant as *better*, but rather as an ordering from first principles, developmental history, and size in Bernal's map.

Ellingham (1948) also uses three columns to orient his map, but the primary axis here goes from top to bottom rather than from left to right. The central column consists of a set of connected disciplines that could be considered more fundamental. From top to bottom we find mathematics, physics, chemistry, biology, and geology. Applied areas are to the right or left of this central column. The left column consists (from top to bottom) of civil engineering, mechanical engineering, chemical engineering, metallurgy, and mining. The right column consists of electrical engineering, chemical engineering, medicine, and agriculture. Social sciences are not included in this map. The columns are of roughly equal size. Therefore, the fundamental sciences (the central column) only cover about one-third of the map.

We consider Ellingham's map as *hierarchical* along one dimension and *noncentric* along the other. The central column represents the traditional ranking of disciplines. The left and right columns emphasize the branches from these central disciplines. Given that they are branching points, physics, chemistry, and biology could be considered as central, which would suggest a centric map. However, the map has overriding noncentric features. In the words of Ellingham (1948, pg. 480),

> In many respects the outer edges of these side panels could properly be joined by wrapping the chart around a cylinder; thus Mechanical Engineering and Electrical Engineering would thereby be justifiably brought together, as well as the two areas which it has been convenient to provide for Chemical Engineering.

Balaban and Klein (2006, Figure 2) present a much more recent expert-based map, and argue that that science is *hierarchically* ordered. The same order of fundamental disciplines is suggested—mathematics, physics, chemistry, and biology—followed by applied areas. Branches from this *central* core deal with the macroscale (earth sciences, environmental science) or the nanoscale (computer technology and engineering), or were branches off of biology (brain and medical science, agricultural science). These branches then converge to a single node at social sciences. This map was unique among all maps examined here in one point: The humanities were placed at both the top (logic feeding into mathematics) and the bottom (law and ethics) of the hierarchy. Thus, although Balaban and Klein argue for a hierarchy of disciplines, it takes little imagination to complete the circle (bottom to top) by linking the two humanities areas.

Bernal (1939), Ellingham (1948), and Balaban (Balaban & Klein, 2006) each stress the hierarchical nature of science. All agree that there is an ordering between mathematics, physics, chemistry, and biology. Medicine might be fifth in this set, but all three maps place medicine as an applied area that is proximate to biology.

Neither Bernal (1939) nor Ellingham (1948) suggest that there is a dominant discipline that is centric. Rather, they suggest that each fundamental discipline has its own set of applied sciences. Balaban and Klein (2006), however, argue that chemistry is more centric. In their map, it is the highest

DOI: 10.1002/asi

science in the hierarchy where branching occurs, and gives rise to four applied areas (earth science, environmental science, computer technology, and engineering). They argue that chemistry is more central than biology since biology is both lower in the hierarchy and only gives rise to three areas (brain science, medical science, and agricultural science).

Reference Paper Maps

The earliest attempts to map all of science using bibliometric techniques were made by Henry Small and his colleagues (Griffith, Small, Stonehill, & Dey, 1974; Small, 1999; Small & Garfield, 1985). These bibliometric techniques focused on highly cocited papers (pairs of references in bibliographies that cooccur perhaps five or more times in one year). In those early days, and due to the high computational costs involved, high thresholds were used, resulting in relatively small samples of documents. These small samples resulted in extreme disciplinary biases. Medicine had the clear advantage since medical papers were more highly cocited than those in other disciplines. Chemistry papers also had reasonably high citation levels, but the remaining sciences (including mathematics and physics) and the applied sciences had much lower citation levels. These relatively different citation rates by discipline persist today.

Small's first map in 1974 illustrates this bias in disciplinary citation levels. The high thresholds used then resulted in the selection of only 1,150 papers to represent all of science. Two nodes (out of 41) dominate. One node (medicine) accounted for 70% of the papers and the second largest node (chemistry) accounted for 8%. Of the 84 edges (links between nodes), 26 connect to the medicine node and 20 connect to the chemistry node. The remaining graph is mostly a dispersed set of nodes that are branches off of medicine or that link to both medicine and chemistry.

This map can be clearly categorized as a *centric* map. However, it is important to emphasize that Small was not suggesting that medicine was the central discipline of science. His training was initially in physics. The cocitation method he developed was conceived of while he was working at the American Physical Society on a project to map the history of physics. This map was a proof of principle, and reflected the inherent biases of using citations when references in medicine tend to be more highly cited.

Small's second map (Small & Garfield, 1985) was the first attempt to overcome disciplinary bias and provide a map that conformed more to common beliefs that physics had a more central role in science. This map was able to replicate the expected set of disciplines that were identified by experts. However, Small de-emphasized the hierarchical nature of science by ordering the disciplines from right to left. On the far right was mathematics (a relatively small node with few branches). This was followed by physics (the second largest node), a set of smaller chemistry nodes, and then one very large node that captured cell biology and medicine. The cell biology/medicine node in this map was once again the largest node.

We have listed this map as a combination of *hierarchical* and *centric*. The expected hierarchical order of disciplines is found, albeit in reverse order and not shown as a linear ordering. The highly centric nature of medicine/cell biology, consistent with the 1974 map, is also shown.

Small's third map (1999) was a further attempt to overcome disciplinary bias. This map is more similar to expert maps. The largest node represents the physical sciences (physics and chemistry are combined, as suggested by Bernal, 1939). The next largest node is biology, followed by medicine and then an area of the social sciences. These nodes are placed in a more traditional ordering, from left to right.

We considered this map as *hierarchical*. There is no single centric node in this map. The grouping and ordering of disciplines are similar to the hierarchical map of Bernal: The physical sciences are predominant, followed by biology, and then the social sciences. The differences between this map and the expert maps are in the size of two applied areas of science, medicine and engineering. The experts only allocate 5% of their maps to medicine versus 20% in this map. The experts allocate almost 50% of their maps to engineering, while engineering can hardly be found on Small's map. These differences are likely due to a combination of actual changes in the distribution of science over 50 years' time (increase in medical research), and disciplinary biases (which decrease the relative share of engineering).

Klavans and Boyack started working together in 2003 in an effort to scale up existing techniques to where millions of papers could be accurately mapped. Both researchers had their training in engineering, and were sensitive to the fact that the applied sciences were still not adequately represented. Klavans and Boyack (2006b) found that disciplinary biases could be significantly reduced by increasing the sample size dramatically, from a few thousands of reference papers to nearly one million reference papers. Using a recursive clustering technique similar to that used by Small (1999), but without excluding references at each subsequent clustering level, the map that emerged (Klavans & Boyack, 2008) had a circular or *noncentric* shape. The same hierarchical ordering of disciplines suggested by experts was found (mathematics, physics, physical chemistry, chemistry, and biochemistry), along with a second sequence linking biochemistry, biology, and medicine. However, this map was circular rather than hierarchical in that the ends of the hierarchy were explicitly linked through the sequence of medicine, psychiatry, psychology, social sciences, computer science, and mathematics.

Klavans and Boyack also explored disciplinary bias as a function of bibliographic database. All literature-based comprehensive science maps created between 1974 and 2006 used the TS citation databases. For 30 years, these were the only databases with sufficient scope, and with sufficiently clean bibliographic information, to be used for this purpose. In 2004, Elsevier introduced a competitive database, Scopus, that claimed to have better representation of the applied areas. Two separate maps of science based on these two databases were generated as a basis for comparing their coverage and

DOI: 10.1002/asi

corresponding impact of disciplinary bias on a map of science (Klavans & Boyack, 2007). The TS-based science map and the Scopus-based science map were very similar in structure and layout. The same circle of science appears with the same general ordering seen in other maps. However, each map also shows earth sciences in a more central position than seen in any other map. Despite this fact, we have classified both of these maps in Table 1 as *noncentric*; in each case earth sciences only links to one side of the ring; thus the ring structure appears to be a more dominant feature than the centrality of earth sciences.

Journal Maps

The next seven maps are based on a different method for partitioning science: using clusters of journals that one might then call *disciplines*. These maps are typically generated in two steps. The first step is to divide journals into some number of clusters, and the second is to generate a layout of the clusters using a layout or visualization algorithm.

Journals have been used as a basic unit for mapping science for some 35 years, starting with the pioneering map of Narin, Carpenter, and Berlt (1972). We do not include this map in our study because it does not meet our measure of being comprehensive. But we note that it was a *hierarchical* map duplicating a portion of the hierarchy mentioned many times above, starting with mathematics, and proceeding through physics, chemistry, and biochemistry to biology.

The first comprehensive journal-level map of which we are aware comes from a research group in France (Bassecoulard & Zitt, 1999). Using a thresholded set of some 2,000 journals, they grappled with questions such as handling of general journals, the choice of a measure of journal:journal relatedness, and clustering or classification method. Although they used a hierarchical clustering method, their map emerges as a combination of the *hierarchical* and *centric* forms. There is strong evidence of hierarchy. Physics is in the upper left. The next major node is engineering, which branches out into chemistry on one side and computer science on the other. The third major node is biology. The centric nature of the map is suggested by the large size, central location, and larger number of links from biology and biochemistry. There is no evidence that the map is noncentric (branches from medicine do not link back to physics via social sciences or computer science).

This map allocates roughly 15% of its area to the engineering disciplines and almost half to the medical fields. Some of the links between fields that we have come to expect from viewing other maps could only be partially observed, for instance between computer science and math; physics, physical chemistry, and chemistry; and biochemistry, medicine, and brain research. Some topics that were expected to be proximate (such as physics and physical chemistry) had intervening nodes. The reasons for these differences are difficult to determine. They may be due to the low sample size (only 2000 journals) or the layout algorithm. It may be possible, but far less likely, that these differences are due to fundamental differences in the structure of science.

The next five journal-level maps were generated by members of the U.S. research team of Klavans, Boyack, and Börner. The first map, by Klavans in 2002 and presented as a poster at the Sackler Colloquium on Mapping Knowledge Domains (Shiffrin & Börner, 2004), appears as a circle (*noncentric*), with roughly the same ordering of disciplines as described previously (mathematics, physics, physical science, chemistry, biochemistry, biology, infectious disease, medicine, brain research, social science, computer science, and connecting back to mathematics). Engineering and health services were not identified as nodes on this map. Historically, among the 20 maps considered in this study, this was the first science map that proposed that the underlying structure of science consisted of a circle of disciplines.

The next journal-level science map was generated with a totally different purpose. Boyack, Klavans, and Börner (2005) were interested in measuring the accuracy of a set of maps of science in order to choose the most accurate relatedness measure for mapping. The TS journal-category structure was used as the standard against which the various maps were compared. Only their most accurate map is reviewed here. This map locates 205 journal clusters, and shows the dominant citation flows between clusters. Although the visualization filled the space (leaving no white space) and there are clusters that appear to be in the center of the map, examination of the linkages shows that there is no central node. Nor is there any evidence of a dominant, linear pathway through the regions of the map. However, one could find the same pathways of linked disciplines associated with the noncentric models described previously. Thus, we labeled this map as *noncentric*.

The third journal map by this research group used two levels of clustering with the VxOrd visualization algorithm, now known as DrL (Martin, Brown, Klavans, & Boyack, 2008). As a result, this map contains far more white space than previous maps, and has been used as the basis for a study of evolution in chemistry research (Boyack et al., 2009). It is *noncentric* in form, and is very similar to the first map by Klavans, despite the use of different data sources and layout algorithms.

The last two journal-level maps by this research team explored the effect of other biases in the TS databases on the shape of these maps. First, Boyack (2009) combined the TS Proceedings database with the Science and Social Science databases. The Proceedings database had extensive coverage in computer science and engineering. Inclusion of these data resulted in a map that placed physics and chemistry in the center of a pentagonal shape. The outer edge of this pentagonal shape was continuously connected, indicating a ring-like structure. Although physics and chemistry are inside the ring, most of their links are to areas at the top (computer science) or right (physical chemistry, chemistry, and engineering) of the map. For these areas on the interior of the map to be considered as centric, there would need to be extensive linking to the lower and left portions of the map. No such links are shown. We thus labeled this map as *noncentric*. We note that, by including the Proceedings database, the applied areas (specifically engineering and computer science) were better

DOI: 10.1002/asi

articulated, and the disciplines feeding these areas (physics and chemistry) had a more centric role in the map.

The most recent journal-level map by Klavans et al. (2008) combined the TS and Scopus databases in order to generate a more complete map in terms of journal coverage. By combining these data, it was possible to generate a matrix indicating the relationship between more than 16,235 journals, proceedings, and book series. This map differs from all of the other maps in that it was laid out on the surface of a sphere. All of the other maps were laid out on a plane, with the exception of the expert map by Ellingham, who proposed that his map be wrapped around a cylinder. Ellingham had actually proposed a spherical layout, but had not implemented it. The reason that a spherical layout was adopted by Klavans et al. (2008) was well articulated by Ellingham (1948, pg. 480) almost sixty years ago:

> By suitable reconstruction it would be possible to allow for the chart to be spread over the surface of a sphere and this would have the advantage of avoiding the need to select a particular science to occupy the centre of the picture.

This map is clearly *noncentric*. The same circle of science, as observed previously using Euclidean projections, can be observed as circumscribing the sphere. Some branching and reconnection does occur. For instance, a string formed by engineering, earth science, and biology branches off from physics and chemistry, and reconnects at biochemistry. Health services branches off from medicine, and reconnects with psychiatry.

The last journal-based map was recently published by researchers at the University of Washington (Rosvall & Bergstrom, 2008). This map, using methodologies that draw more from network science than from bibliometrics, appears as a circle of science that is similar to many of the *noncentric* maps listed in Table 1. There is the expected sequence of mathematics, physics, physical chemistry, chemistry, biochemistry, infectious disease, medicine, brain research, psychology, social science, computer science, and back to math. Engineering areas are also noted, but are not linked together.

Journal Category (Discipline) Maps

The next three maps use the TS disciplinary classification system. TS has two different sets of journal categories, a broad set with some 25 categories, and a finer-grained set containing well over 200 categories. At the fine-grained level, journals are assigned to one or multiple categories. The average number of categories per journal is 1.6 (Leydesdorff & Rafols, 2007), thus providing the overlap to enable co-category types of analyses.

Two of the discipline-based maps in Table 1 were created by the SCImago research group in Spain. Their two maps, however, show vastly different and contradictory pictures of how science is structured. The first map (Moya-Anegón et al., 2004), which used the broad category structure, replicates the circle (*noncentric*) of science mentioned previously, and shows the expected linkages mentioned above in conjunction with other noncentric maps.

The second map used the fine-grained category structure, and thus has far more detail (Moya-Anegón et al., 2007). However, instead of showing a circular structure, this map is highly *centric*. This map is the very epitome of a hub-and-spokes type of diagram. The largest TS journal category, Biochemistry and Molecular Biology, is the central node; all other nodes attach back to the central node through one of the 30 or more separate branches. The largest branch (47 nodes) leads to biology, and then branches out to mathematics and the social sciences. The second largest branch (43 nodes) leads to chemistry, engineering, and physics. The third, fourth, and fifth largest branches lead to neuroscience, pharmacology and earth sciences, and general medicine and health services, respectively. This is the only map reviewed here that uses Pathfinder networks (PFNet) for layout. In most cases, PFNet constrains the number of links to be one less than the number of nodes. This extreme pruning of links between categories leads to disconnection of links that are found in the majority of the other maps, and implies that mathematics is not linked to computer science, social sciences are not linked to psychology, and chemistry is not linked to geochemistry. This extreme pruning of links creates a map with features that are nonintuitive and in disagreement with most every other map listed in Table 1.

Another discipline-based map was recently generated by Leydesdorff and Rafols (2007). We consider this to be a *mixed* map, much like the journal map by Bassecoulard and Zitt (1999), in that it contains examples of multiple forms. There is evidence of the expected hierarchical split between the physical sciences (physics and chemistry) and the biological sciences (biology, biochemistry, and medicine) that was proposed by Bernal (1939). There is evidence of noncentricity: The physical and biological sciences are linked by two possible pathways: via the geosciences or via the computer sciences. This map does have an intriguing feature: There is a major link between biochemistry and computer science that is unique among the maps reviewed here, a link that captures the importance of bioinformatics.

Other Maps

The final map in our review presents a totally different approach to mapping science. Balaban and Klein (2006, Figure 3), in the second map shown in their paper, suggest a hierarchy of science based on examination of course prerequisites in the undergraduate catalog of courses offered at Texas A&M University. Their approach to partitioning is similar to the one used by experts (including their own "expert" map from the same paper); they start with traditional categories and then modify the categories based on the collection of course requirements. Their approach to linkage and relative location in the hierarchy is based on prerequisites (which courses must be taken first). This rather unique approach results in their hypothesis that chemistry is the central discipline in science. Their map is therefore labeled as *centric*, even though the authors present a picture that

DOI: 10.1002/asi

looks noncentric—they show that mathematics, at the top, is directly linked to the bottom discipline of social sciences.

Data Biases and Map Forms

This diverse set of maps, generated using a variety of datasets, methods, and algorithms, provides strikingly similar results in terms of proximity of pairs of disciplines, but also varies widely in terms of form (hierarchical, centric, noncentric, or mixed). Although this variance in form can result from differences and biases in data sources, as well as differences in choices of relatedness measure, visualization algorithms, and edge pruning, we find that form does tend to correlate with biases in the input data sets.

Maps based on the least biased data sets (by this we mean maps with the most comprehensive coverage in terms of their input) that are not dominated by a few extremely large nodes tend to be generated in the noncentric form. Maps based on data that has been thresholded, or that are dominated by one or a few highly dominant nodes, tend to be generated as hierarchical or centric maps. Bias can be introduced in several ways: through database bias, by setting arbitrarily high thresholds for the data, or by selecting an inappropriate measure of relatedness.

Database bias is particularly evident in the comparison of the TS Science Citation Index and the Scopus database. The TS Science Citation Index is weighted somewhat toward life sciences and medicine. The Scopus database includes the majority of journals covered by TS, but adds a significant number of journals and proceedings from engineering, computer sciences, and health services. The TS Proceedings database, when added to the TS Science Citation Index, compensates somewhat for this bias (Boyack, 2009). When the TS and Scopus databases are combined, the resulting map is even more balanced, and a noncentric map emerges.

High thresholds also introduce bias, and tend to result in maps that are either hierarchical or centric in form. Computer science, social sciences, and the humanities, when poorly represented or not represented at all, tend to disconnect the circular shape into a linear form that is then interpreted as hierarchical. This is especially apparent in the two maps of Balaban & Klein (2006). Their expert-based map clearly shows the underpinnings of humanities at the top and bottom of the hierarchy (logic is above math; ethics is below the social sciences), but fails to link the two. Their course-based map clearly states that the social sciences are at both the top and bottom of the hierarchy, but again fails to complete the circle. In other words, both of these maps could be interpreted as noncentric, even though they are originally shown as hierarchical. Another example of this effect can be seen in the recent journal map by Samoylenko, Chau, Liu, and Chen (2006). This map links journals above an impact factor of 5 using minimum spanning trees. Given this high threshold, mathematics, computer science, and engineering are entirely excluded. We thus do not consider this map as comprehensive due to its exclusion of fields, and have chosen not to include it in our review and comparison.

The choice of the relatedness measure can also significantly bias the map. Boyack, Klavans and Börner (2005) measured the accuracy of eight different maps, each generated using a different measure of journal:journal relatedness. They found that the least accurate map was based on using raw co-occurrence frequencies. One can also see in their paper that the least accurate map, generated from raw co-citation counts, is centric in form (pg. 359, Figure 1, lower left). We are thus not surprised that the recent SCImago map (Moya-Anegón et al., 2007), which was generated from raw category:category co-citation frequencies (modified by small additions to avoid duplicate matrix entries), is highly centric. We cannot assume that a map generated from raw category:category co-citation counts is as inaccurate as a map based on raw journal:journal co-citation counts without testing. However, the structural similarities between the two maps suggest that the accuracies cannot be too far different. It would be very interesting to see if the same map generated from a more accurate relatedness measure would produce a centric map. We expect that it would not.

A Consensus Map of Science

As mentioned in the section on selection criteria, to be decomposable, maps of science must have partitions and links. In this section, we introduce a consensus map of science that is based on 16 partitions in science. We then examine the proximate locations of these 16 partitions of science on the 20 maps of science from Table 1. A link between pairs of partitions is counted as a consensus link if more than half of the maps advocate that particular link. The correspondence of this consensus map with the 20 input maps is then examined from multiple perspectives. Possible shortcomings of the consensus map are discussed.

Areas of Science

We divided science into 16 broad areas for purposes of codifying the 20 input maps (see Table 2). We started with the four fundamental areas mentioned in most maps (mathematics, physics, chemistry, and biology) and then considered the six possible combinations of these four areas. Only two of the combinations (physical chemistry and biochemistry) were found to occur with any frequency among the 20 input maps. We then identified another six areas that were more applied. Three areas (computer science, engineering, and geoscience) represent the applied areas building off of

TABLE 2. Sixteen areas of science used to characterize the 20 input maps of science.

M – Mathematics	B – Biology
CS – Computer science	I – Infectious disease
P – Physics	MD – Medical specialties
PC – Physical chemistry	HS – Health services
C – Chemistry	N – Brain research (neuroscience)
E – Engineering	PS – Psychology/psychiatry
G – Earth sciences (geoscience)	SS – Social sciences
BC – Biochemistry	H – Humanities

DOI: 10.1002/asi

mathematics, physics, and chemistry. Note that we explicitly include electrical engineering with computer science, while the engineering area is comprised of all engineering disciplines other than electrical engineering. The other three areas (infectious disease, medical specialties, and brain research) represent the applied medical-related areas related to biology.

An additional three areas (health services, psychology, and social sciences) represent applied areas that deal more with social issues than with the hard sciences. These areas are very large and diverse fields that were not well represented in the expert maps or the bibliographic maps using the TS database. The addition of the Scopus database helps to better represent the role of these applied areas in science, particularly in the case of health services (which includes nursing). The final area, humanities, could be considered fundamental to the social sciences (Balaban & Klein, 2006; Bernal, 1939). Unfortunately, only a few of the maps in Table 1 explicitly locate this area of research. However, given that one citation database is specifically geared to the humanities, the TS Arts and Humanities Citation Index, we felt it best to explicitly include it as a separate area. Scopus has very scant coverage of the humanities.

Although we realize that there is a certain subjective nature to the selection of these 16 areas, and that other researchers might define a different set of areas, we find a reasonable balance in using the areas shown in Table 2. There are seven fundamental areas and nine applied areas. The nine applied areas are equally split between the physical sciences, the biological sciences, and the social sciences. The overall balance between the 16 areas can be seen by examination of the relative sizes of different regions in the most recent journal map by Klavans et al. (2008), which has the broadest coverage of any map to date.

Coding of Input Maps

Each of the 20 maps of science reviewed here was analyzed in detail to determine the locations of the 16 areas, overlaps or proximate locations of pairs of areas, and additional linkages between areas that were not proximate. This was done using a four-step process as follows.

First, the 16 areas of science were located on the 20 maps of science. In many cases this was done by simply placing a single node for an area at the location on the map labeled with the area name. In cases where a map was extremely complex (many nodes and edges), the dominant locations of the 16 areas were found, and nodes placed at those locations. There were some cases where an area seemed to be located in multiple positions on one map; if so, this feature was preserved by locating multiple nodes for that area. On occasion, a node was labeled with multiple areas if it was clear that the author was referring to a broader area of science than indicated by only one of the areas listed in Table 2. This was also done if the intentions of the author were unclear, although this was rare. Note that not all maps contained all 16 areas of science.

Second, links (or edges) between areas of science were drawn if the map and text suggested that these areas had overlapping domains (proximate location) or were otherwise connected (linkage). In cases where an input map showed many nodes and edges, an edge was drawn between two areas of science if the sum total of the original nodes and edges seemed to indicate a strong relationship between the two areas. Thus, only the dominant relationships between the 16 areas were captured if the edges in the map had not been pruned.

At this point in the process, each of the 20 input maps had been simplified into maps comprised of only the 16 (or fewer) areas of science from Table 2 along with the dominant linkages between areas. In each case, the simplified area map was actually overlaid on top of the input map. Our third step was to further simplify these maps. A map was simplified if (a) two nodes with the same area code were linked (such as an engineering area linked to another engineering area), or (b) an edge was redundant. For example, if a medical node linked to two different brain-research nodes, it was simplified to a medical node linked to one brain-research node. The final step in the coding of each map was to convert the map to a set of links based on paired relationships between areas. This set of links for each map was then used as the basis for further analysis.

To further clarify the coding process, we provide an illustration (see Figure 2) of how one of the maps was coded. We illustrate the process with the SCImago-II map (Moya-Anegón et al., 2007) because it illustrates many of the issues about biases that have been previously raised. The first frame in Figure 2 is a copy of the map as originally published. The second frame overlays the 16 areas of science listed in Table 2 using Steps 1 and 2 of the coding process detailed above. Note that the central node in the coded map, which represents the central node and several of the immediately surrounding, but singly linked nodes in the SCImago-II map, is given three area assignments (BC–Biochemistry, MD–Medical Specialties, and I–Infectious Disease).

The major branches of the network in Figure 2b are represented using nodes and edges. The network on the right involves the disciplines normally associated with the hierarchy of science: (C) chemistry, (PC) physical chemistry, and (P) physics. Different (E) engineering disciplines branch off of chemistry and physical chemistry, and the (CS) computer science discipline branches off of physics. Note, however, that mathematics is not found anywhere near this network. Mathematics is on the opposite end of the map, emerging from the central node through (B) biology to (M) mathematics and then ending in the (SS) social sciences, (H) humanities, and (E) engineering.

The third frame (Figure 2c) simplifies the network using the rules in Step 3 above, and marks those areas that have duplicate locations in gray. There are a large number of duplicate areas on this map. Engineering is in four different locations, while the social sciences and geoscience each have two separate locations. There are far more instances of multiple locations in this map than in any of the other input maps. We believe this to be an artifact of the relatedness measure and edge-pruning algorithm used to generate this map.

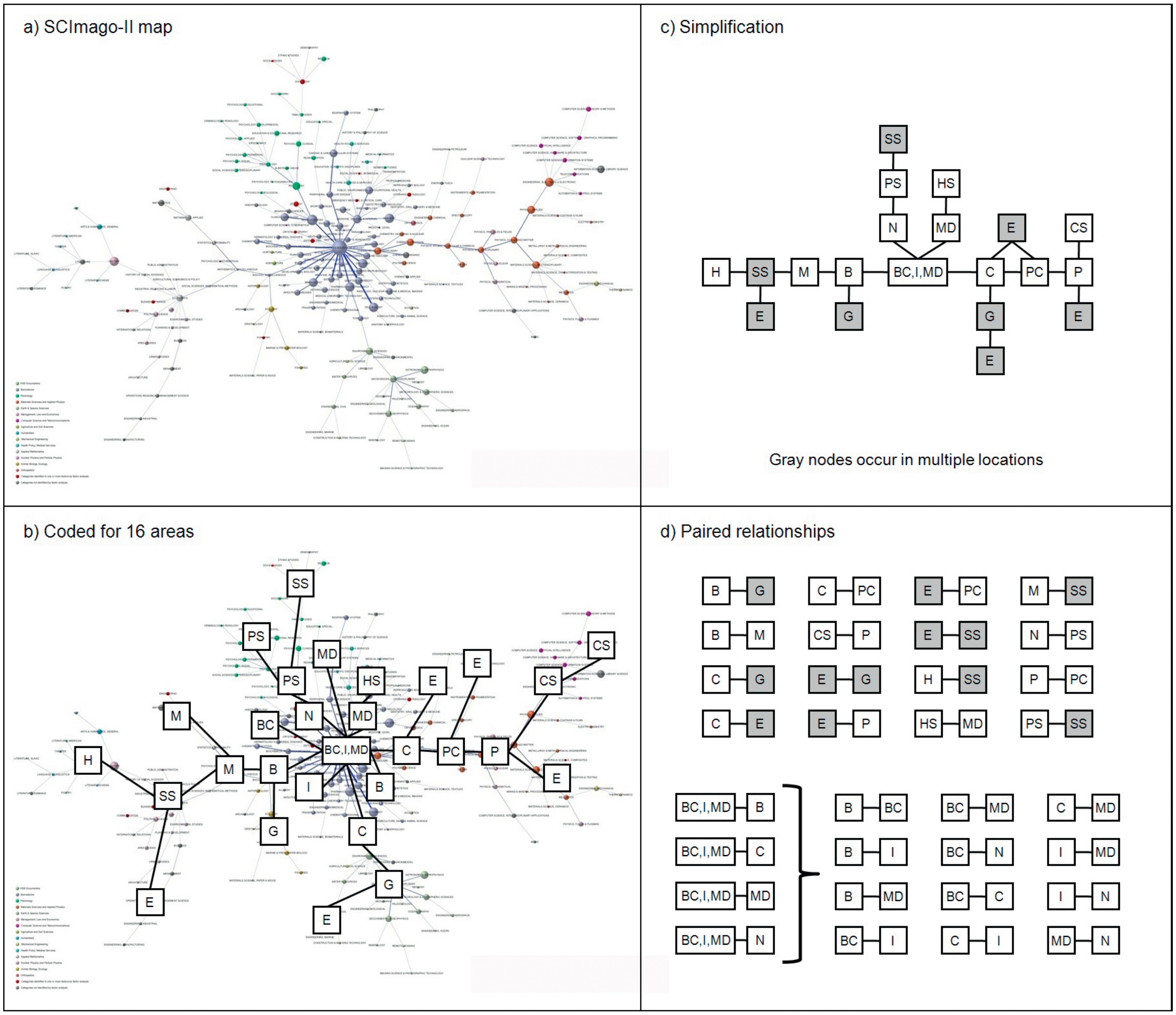

FIG. 2. Coding of the SCImago-II map (Moya-Anegón et al., 2007). The other 19 maps were all coded in a similar fashion.

When raw-count relatedness measures that are dependent not only on discipline size but also on discipline citation culture are used, smaller, lower-citing disciplines that are part of one larger area can end up dispersed to far-flung locations. This happens because these smaller disciplines link preferentially to much larger disciplines in different scientific areas due to the overall weight of the raw counts from the larger discipline, rather than linking together in intuitive ways.

The fourth frame (Figure 2d) shows the conversion of the simplified map into linked pairs of areas. There are 20 edges in the network shown in Figure 2c. Note that four of these edges link to the central node, which has three area assignments. These edges are shown at the lower left of Figure 2d as four coding pairs. For purposes of generating a consensus map of science, we expand these four coding pairs into all of their unique permutations (including the ones inside the triple node, BC-I, BC-MD, and I-MD), which are shown as the 12 pairs at the lower right of Figure 2d. In total there are 28 pairs of linked areas, 16 of which come directly from edges, and another 12 that come from permutations from edges associated with multiarea nodes. Note that some paired relationships that one would expect to see (such as mathematics being linked to either computer science or physics) do not appear in this representation. This would be one example of local inaccuracy if there is a consensus that mathematics is linked to physics.

Codings for all 20 maps are shown in Figure 3 in a manner similar to that shown in Figure 2b. In addition, high-resolution images of all 20 original maps and the codings for those maps are available online at www.mapofscience.com/history/maps. In the following sections, all paired relationships from all 20 maps are analyzed in order to establish a group consensus and the relative correspondences of each of the input maps.

Consensus Maps of Science

One-dimensional and two-dimensional consensus maps of science have been generated from the dominant paired

DOI: 10.1002/asi

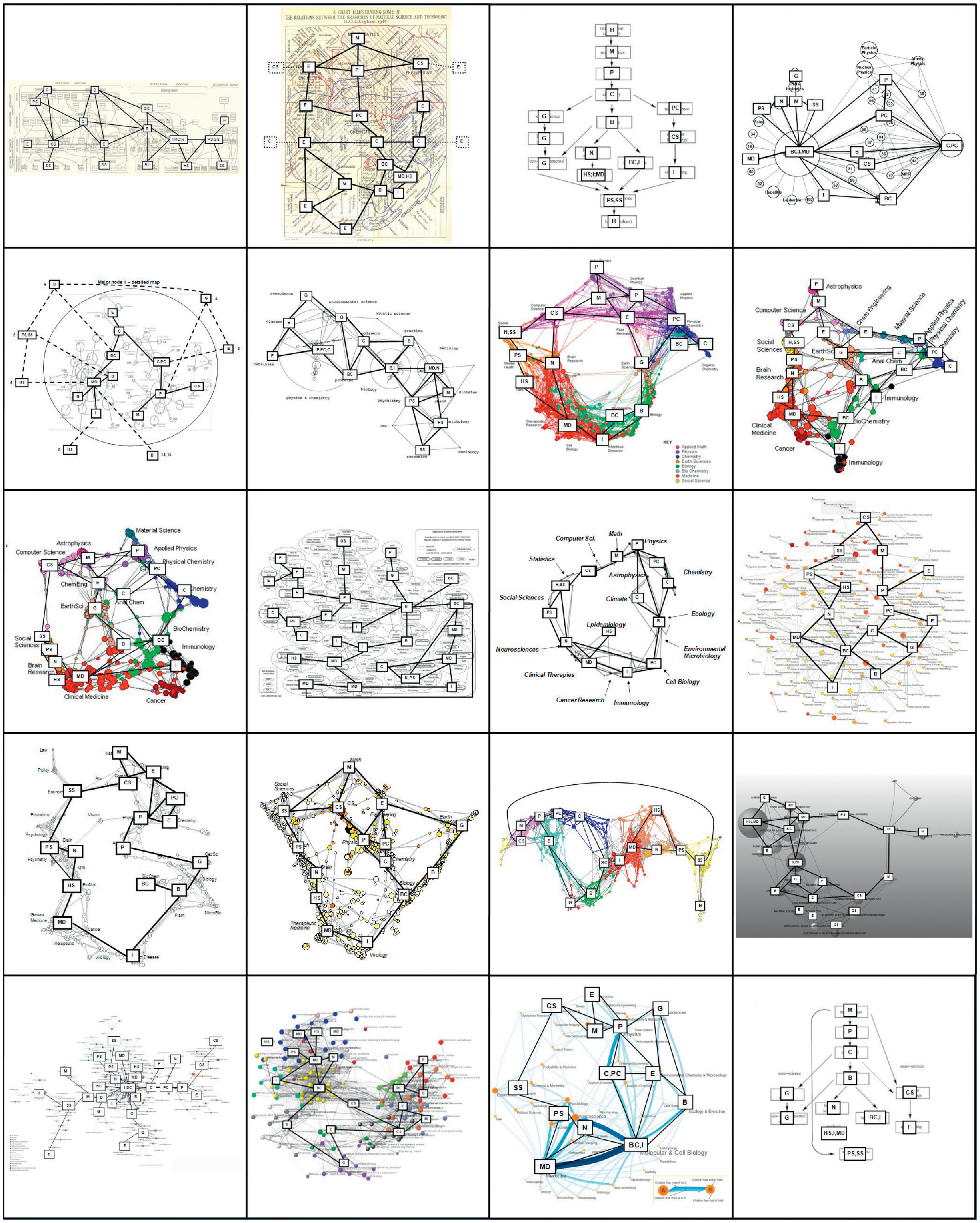

FIG. 3. Images of the 20 maps of science that were used in this study along with their codings. The 20 maps are shown in the same order in which they are listed in Table 1, from upper left to lower right.

relationships between areas across the 20 input maps of science. First, all paired relationships from each of the input maps were coded into a database. A *consensus link* was then defined as any paired relationship that occurred in at least 50% of the maps in which it could have occurred.

We started with the initial list of all edges (400 in total) as given in Appendix A. We then generated all of the permutations between pairs of areas arising from the implied edges when a single node had multiple area assignments (e.g., the central node in Figure 2a with assignments to BC, I, and

DOI: 10.1002/asi

TABLE 3. Consensus pairs of scientific areas from 20 maps of science.

Rank	Pair	N	N-poss	%
1	B-BC	20	20	100.0
2	I-MD	20	20	100.0
3	H-SS	8	8	100.0
4	C-PC	19	20	95.0
5	HS-MD	16	17	94.1
6	PS-SS	16	17	94.1
7	P-PC	18	20	90.0
8	MD-N	16	18	88.9
9	E-G	16	18	88.9
10	B-G	17	20	85.0
11	BC-I	16	20	80.0
12	E-PC	14	18	77.8
13	N-PS	14	18	77.8
14	CS-M	13	18	72.2
15	BC-MD	14	20	70.0
16	BC-C	14	20	70.0
17	E-P	12	18	66.7
18	B-I	13	20	65.0
19	CS-SS	10	16	62.5
20	H-PS	5	8	62.5
21	M-P	11	19	57.9
22	C-E	10	18	55.6
23	C-P	11	20	55.0
24	HS-N	8	15	53.3
25	CS-E	9	17	52.9
26	C-G	10	20	50.0
27	HS-PS	8	16	50.0

MD). This brought the total number of paired relationships to 535. Less than 10% of the nodes (26 out of 275) in all 20 maps had multiple area assignments. Only seven nodes had three area assignments, and none had four or more area assignments. We then calculated which pairs of relationships could appear in which maps. Some pairs of areas could not occur in all 20 maps. For example, the engineering (E) and physical chemistry (PC) areas were only present together in 18 of the 20 maps. They were connected by an edge in 14 of those 18 maps; thus this edge occurred 77.8% of the times possible. Consensus links are listed in Table 3.

Two-Dimensional Consensus Maps

Figure 4 illustrates how the pruning of edges affects the final layout of a two-dimensional map generated from consensus pairs of scientific areas. Figure 4a shows the effect of extreme edge pruning. In this map, only the top 15 edges (and the tie for # 15) from Table 3 were used to generate the map. A hierarchical picture emerges. One sees a similar ordering of areas to that seen in many of the individual maps; the hierarchy starts with physics and continues through engineering, the earth sciences, and biology/biochemistry. There is also a separate branch from physical chemistry through chemistry to biochemistry. The hierarchy then continues through the medical areas to the social sciences and humanities. Computer science and mathematics form a separate component in this map; they do not link to the large component at a high level of edge pruning.

Figure 4b shows what the two-dimensional consensus map looks like if all of the edges from Table 3 are included. In this case, we get the noncentric form, with the familiar progression of scientific areas seen in so many of the individual input maps. Mathematics, arbitrarily placed at the top, is followed clockwise around the circle by physics, chemistry, biochemistry, and biology, the medical areas, brain research, psychology, social sciences, and computer science, ending up back at mathematics. Small branches off this main circle pick up the other areas of engineering, geology, and the humanities. In neither case can one duplicate the centric pictures suggested by some of the input maps in Table 1. As previously mentioned, we suggest that the centric maps are an artifact of the underlying bias in the database, the use of inaccurate measures of relatedness, and/or a layout that does extreme edge cutting. These biases lead to a false impression that there is a center to science, and should be avoided.

We realize that generation of a consensus map of science in which 8 of the 20 input maps come from members of our collaborative team (Klavans, Boyack, and/or Börner) would lead some researchers to suspect a bias toward our previous results. To alleviate these concerns we have performed the same analysis while excluding all 8 of our input maps. The consensus maps in Figure 5 were generated from the remaining 12 maps using the paired relationships and the same thresholds, and thus have no direct input from maps generated by our team.

Figure 5a shows the hierarchical structure from extreme edge pruning, in which the map has been generated from the top 15 edges. This map has some similarities and some differences from the structure shown in Figure 4a. The biggest difference between this map and the one generated from all 20 input maps is that the single large hierarchical component from Figure 4a is not preserved; it splits into two components: one for physics and chemistry, and one for the balance of the hierarchy. There is also more linking between the life and medical sciences in the lower component than was found in the hierarchical structure of Figure 4a. In addition, both mathematics and computer science are now isolates. We also show a dashed edge between physical chemistry and engineering, which would have appeared if the threshold were lowered by just one edge (this was the 16th edge). Figure 5b shows a similar noncentric shape to that found in Figure 4b. Mathematics and computer science have switched positions. In general, the number of links within the physical sciences and the number of links within the medical and social sciences have increased. However, the edge between mathematics and physics is no longer present. This map seems every bit as robust as the map in Figure 4b; there are similar numbers of edges in the 90%–100% and 70%–90% consensus ranges in both maps. Comparison of the number of times each of the consensus edges occur in the two maps (Figures 4b and 5b) gives a correlation of .81.

The consensus maps obtained by excluding the 8 input maps from our research team are sufficiently similar to those obtained using the full set of 20 maps that no bias based on dominance of a single research team can be claimed. We thus

DOI: 10.1002/asi

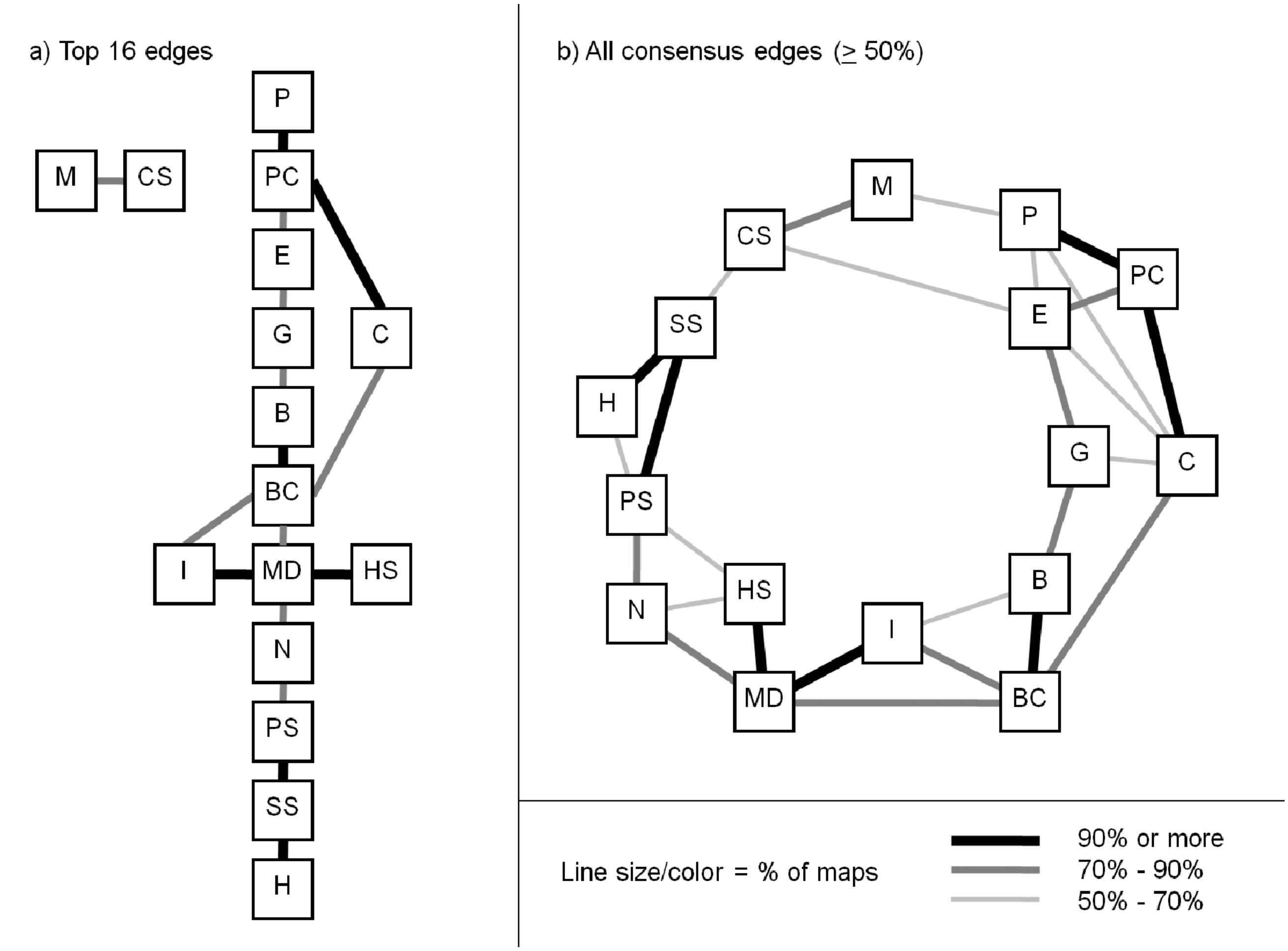

FIG. 4. Two dimensional consensus maps of science from all 20 input maps.

advocate that the noncentric map of science based on all 20 input maps using all consensus edges (from Figure 4b) be adopted as a consensus map of science. We have shown that the map would take a hierarchical form if one wanted to use high thresholds (15 edges) and ignore important relationships. We have also shown that the data does not lend itself to a centric form for reasons given above. The data does, however, generate a noncentric form if one wants to capture the majority of the information from the 20 maps in Table 1.

One-Dimensional Consensus Maps

Figure 4 suggests a solution that is very close to being one-dimensional. If math and physics are linked, 13 of the 16 areas in Figure 4a can be placed in a strict order. Most of the remaining 3 areas can be placed in order so that they are only one hop away from the area they were linked to. Figure 4b, however, does not seem to be as easily collapsed into a circle onto which the 16 areas are consecutively ordered. Despite this, we have calculated the one-dimensional solutions that have the highest correspondence with the 20 existing maps (see Figure 6).

We find that the circular solution has higher correspondence (74.1%) that the linear solution (70.4%). This is not a surprising finding. One can prove, relatively easily, that a solution based on a curved surface (Riemannian space) will be equal to or superior than a solution based on a flat surface (Euclidean space).

Take, for example, the Euclidean solution in Figure 6. If the two ends of the solution are linked, creating the Riemannian solution, the correspondence increases because of the link between CS:SS and the one-hop links between SS:M and CS:H. The Riemannian solution will always be equal to or better than the Euclidean solution because votes for nodes on the outer edges will be added in where applicable. Use of Riemannian space does not mean that the ends must be linked, but rather enables an additional linkage to be made where appropriate.

Correspondence Between Maps

In much of our previous work we have been very careful to establish the accuracy of our methods and the resulting maps (Boyack et al., 2005; Klavans & Boyack, 2006a, 2006b). In this case there is no objective standard that can be used to measure the accuracy of the consensus map, but we can measure the correspondence between the consensus map and the 20 input maps. Multiple aspects of correspondence are

DOI: 10.1002/asi

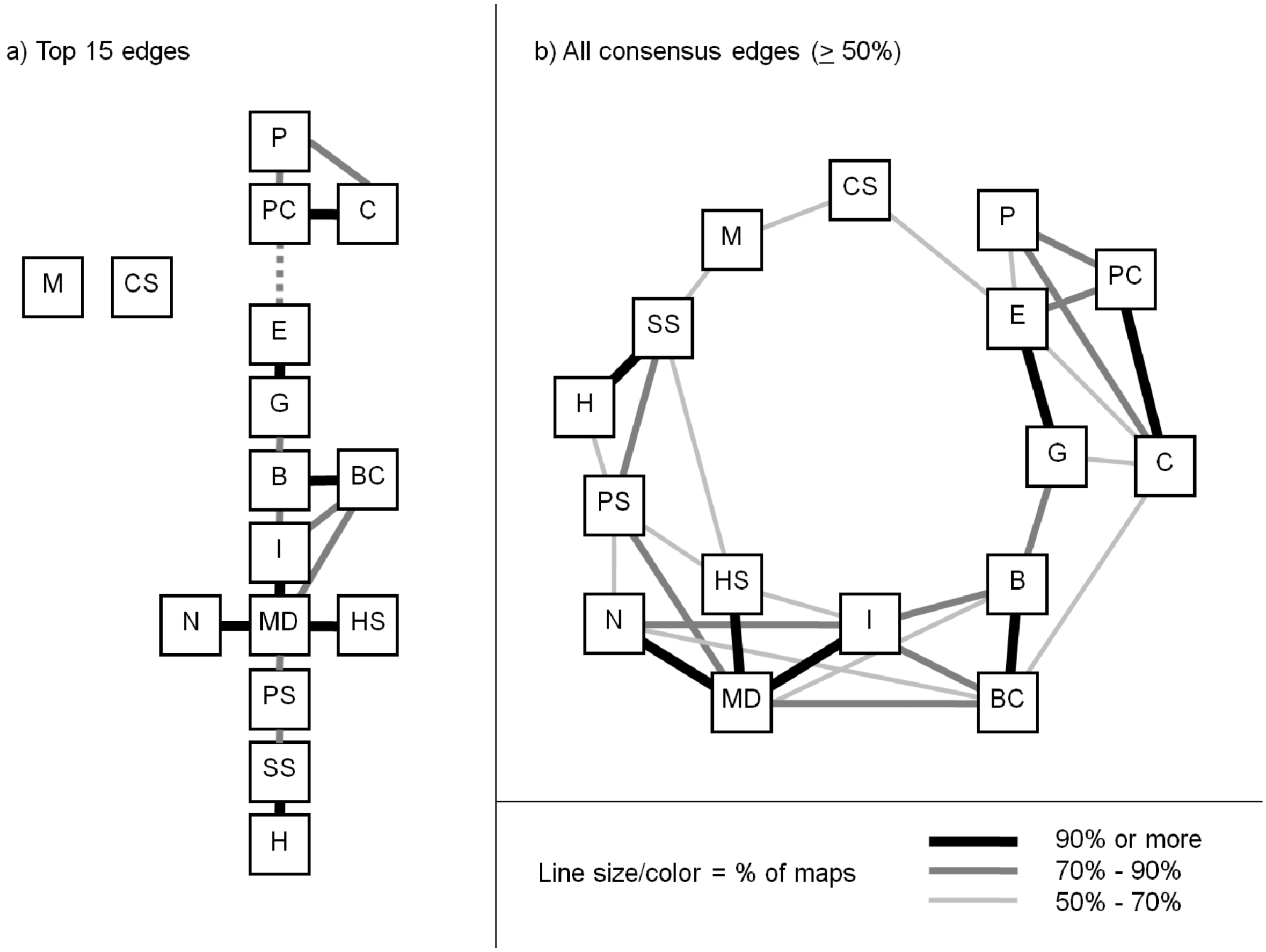

FIG. 5. Two-dimensional consensus maps of science from 12 of the 20 input maps, excluding the 8 input maps from Klavans, Boyack, and Börner.

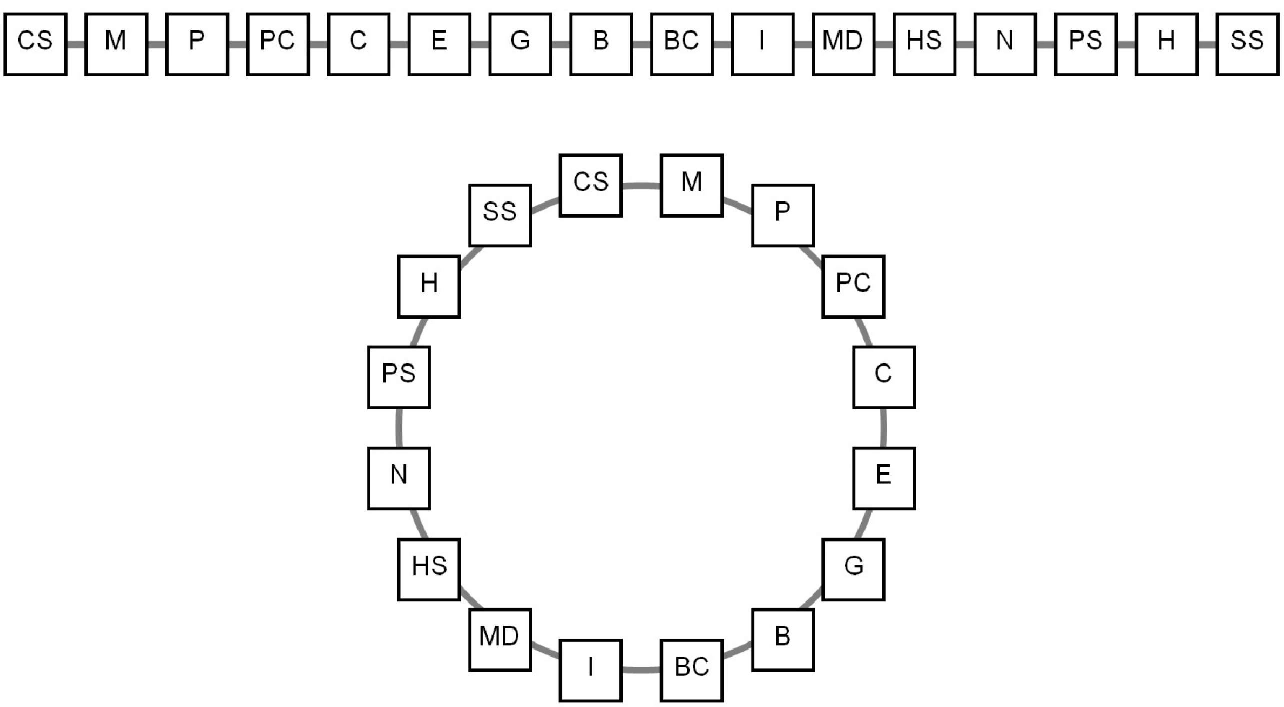

FIG. 6. One-dimensional consensus maps of science, Euclidean (top) and Riemannian (bottom).

TABLE 4. Characteristics of edges from 20 maps of science on the consensus map of science.

Number of hops	Count	Percent	Accuracy value
1	345	78.4	1.0
2	73	16.6	0.5
3	18	4.1	0
4	4	0.9	0
Total	440	100	0.867

examined here. First, we look at the overall ability of the consensus map to capture the data from the 20 maps. Second, we look at the correspondence of each of the 20 maps from three perspectives. Two perspectives focus on local correspondence based on an analysis of paired relationships. This follows the method of Klavans and Boyack (2006a), in which paired relationships in maps are compared with the paired relationships of a gold standard. The third perspective focuses on a measure of regional correspondence, where regional correspondence refers to the ability to put all of the nodes representing one area in the same region of a map (Klavans & Boyack, 2006b). Maps that split up an area of science (such as chemistry appearing on the left and then again on the right) have lower regional correspondence.

Table 4 shows the ability of the consensus map to capture the information in all 20 maps of science. Using the 400 edges listed in Appendix A and the 40 edges inferred by the relationships inside the multiarea nodes (e.g., for the node C;PC, we infer the edge C-PC), we used the consensus map in Figure 4b to determine how many hops between nodes it would take to traverse the path suggested by each of the edges. For example, the simplified SCImago-II map in Figure 1c suggests that biology (B) is linked to chemistry (C). The consensus map suggests that one has make two hops, or traverse two edges (B to BC, and then BC to C), to go from B to C. Each edge was analyzed in this fashion. For edges associated with multiarea nodes, the shortest path from any of the areas in the multiarea node was used. For example, if there was an edge between nodes I;MD and BC;C, the edge was considered to have only one hop on the consensus map if any of the four possible edge combinations (BC-I, BC-MD, C-I, or C-MD) existed on the consensus map. The reason for calculating correspondence in this manner is that we did not want to arbitrarily penalize a map for having multiarea nodes. Over three-quarters of the paired relationships from all 20 maps appear as a paired relationship in the consensus map (number of edges traversed was 1). These 345 relationships were coded as having 100% local correspondence.

We also assumed that traversing across two edges does not indicate a complete lack of correspondence. This is analogous to talking to someone sitting two chairs away while sitting at a large table; it is not necessarily easy to do, but not too hard either. This occurred in 17% of the cases, which we coded as having 50% (partial) correspondence. By contrast, talking to someone sitting three or more chairs away is very difficult. Thus, the 22 cases where one had to traverse three or more edges were considered to have no correspondence. The overall ability of the consensus map to reflect the combined input of the 20 individual maps is 86.7%.

The correspondence of individual maps can be calculated in two different ways. First, one can assume that the consensus map is the gold standard, and count the number of hops associated with each of the edges in a particular source map. This is the method used in Table 4, aggregated to all 20 input maps. Or, one can assume that the individual map is the gold standard, and count the number of hops associated with each of the edges of the consensus map. We calculate correspondences using both of these bases.

Table 5 lists the correspondence of the 20 maps in Table 1 from these two perspectives. Type 1 local correspondence is the latter case, where the individual map (called "Source map" in the table) is the gold standard. This measures how well the consensus map agrees with the source map. Type 2 local correspondence is the former case, in which the consensus map is considered the gold standard. This measure shows how well the source map agrees with the consensus map. In each case, we calculated accuracies using the same method (numbers of hops on the network) and the same correspondence coding assumptions used to calculate the overall correspondence of the consensus map in Table 4. These two types of local correspondence are highly correlated, but not identical. The majority of the difference comes from the denominator of the calculation; in the Type 1 case the number of edges in the source map is used as the denominator, while in the Type 2 case the number of edges (27) in the consensus map is used. Thus, there are cases where a source map is communicating something unique that is not captured by the consensus map. These unique contributions will be discussed in a subsequent section.

Table 5 also lists several other values. Regional correspondence is defined as the ability of a map to put all of the research, for a specific area of science, in proximate locations on a map (Klavans & Boyack, 2006b). In this paper, we defined regional correspondence as the number of unique areas of science represented on the map divided by the number of total nodes. As mentioned above, regional correspondence decreases when a scientific area appears in multiple locations on a map. We note that experts who generated detailed maps by hand (Bernal, 1939, and Ellingham, 1948) made great efforts to maintain a high level of regional correspondence. The map with the lowest regional correspondence was the SCImago-II map (Figure 1). All 16 unique areas of science were represented in their network; however, many of these areas were given multiple positions in dispersed areas of the network. This type of dispersion only occurred in 5 of the 20 maps.

Table 5 is ordered by a figure of merit that is the mean of the three correspondence measures. This ordering suggests that the correspondence of the expert maps has only recently been matched by that of algorithmically generated maps. The two detailed expert maps drawn 60 and 70 years ago by Bernal (1939) and Ellingham (1948) are ranked fourth and sixth in the set of 20 maps, both with figures of merit greater than 93 as

DOI: 10.1002/asi

TABLE 5. Correspondence measures for 20 maps of science.

Source map	Year	Type	Local correspondence Type1	Local2 correspondence Type2	Regional correspondence	Figure of merit	# Areas	Multi-nodes
KB06-SC	2006	Paper	95.8	94.0	100.0	96.6	15	0
Backbone	2004	Jnl	97.6	88.0	100.0	95.2	15	0
UCSD	2007	Jnl	95.7	88.9	100.0	94.8	16	0
Ellingham	1948	Expert	90.0	92.1	100.0	94.0	12	1
KB-Para	2005	Paper	92.3	94.4	93.8	93.5	16	1
Bernal	1939	Expert	85.7	94.0	100.0	93.2	15	2
Scimago-I	2004	Categ	90.9	87.5	100.0	92.8	15	2
KB06-TS	2006	Paper	91.7	90.7	93.8	92.1	16	1
B03-ST	2005	Jnl	92.5	82.0	100.0	91.5	15	0
BBK02-S	2004	Jnl	92.5	80.0	100.0	90.8	15	0
Rosvall	2007	Jnl	78.3	93.2	100.0	90.5	14	2
Small99	1999	Paper	78.6	89.5	100.0	89.3	13	3
Balaban-II	2007	Pre-req	85.0	82.0	100.0	89.0	15	4
K02	2002	Jnl	84.2	81.8	100.0	88.7	15	1
L-R	2007	Categ	86.1	73.9	100.0	86.7	14	0
Balaban-I	2007	Expert	73.9	79.6	100.0	84.5	16	3
Small85	1985	Paper	84.2	76.0	86.7	82.3	15	2
Small74	1974	Paper	69.2	76.5	100.0	81.9	13	2
B-Z	1999	Jnl	80.6	71.7	93.3	81.9	14	1
Scimago-II	2007	Categ	90.0	75.9	75.0	80.3	16	1

evaluated against the consensus map. Although neither map covered all 16 areas of science as listed in Table 2, when compared against the consensus for the areas of science they covered, they agree very well with the consensus. The map with the highest overall figure of merit is the paper-based map generated from Scopus data by Klavans and Boyack (2007). We also note that all of the journal- and paper-based maps generated since 2002 have figures of merit above 88. Paper- and journal-based maps generated before that time were subject to much higher levels of disciplinary bias, and thus have lower figures of merit. Three of the four lowest ranked maps are also those with the lowest regional correspondence values.

It is also interesting that, in general, the number of multiarea nodes increases as we progress down the list. Five of the six maps with no multinodes are ranked in the top 10 maps. This finding was somewhat counterintuitive to us for the following reason. Given the way correspondence was calculated, with edges associated with multiarea nodes getting the highest correspondence value found from any of their permutations, one would expect the presence of multiarea nodes to increase the correspondence value. We thus expected the rankings to show more multiarea node maps near the top. Table 5 shows that this was not the case. The results in Table 5 also suggest that, of the algorithmically generated map types, the paper- and journal-based maps seem comparable, and both have slightly higher correspondences than category-based maps.

Shortcomings of the Consensus Map

The major shortcomings of the consensus map arise from edges (from the 20 input maps) that would have to go part or all of the way across the middle of the consensus map. While this only accounts for 5% of the edges (those requiring three or more hops from Table 4) across all maps, these instances are worth noting. The largest shortcoming of the map concerns edges from mathematics and computer science. Twelve of the 21 edges requiring 3 hops or more go from either mathematics or computer science to more distant nodes across the map. These two disciplines seem to have broad application to many areas. However, due to the low coverage of computer science in most databases, and the lower level of citation in both mathematics and computer science, they tend to be poorly represented in many maps.

The best example of a map that overcomes this shortcoming is the one by Leydesdorff and Rafols (2008). While this map is listed in the lower half of the correspondence listings in Table 5, its lower score is largely due to computer science being located as an important bridge between the physical and biological sciences. Therefore, one might argue that the map does not have lower correspondence, but rather that it highlights the interdisciplinary nature of computer science, which is not picked up as well by other maps.

Another shortcoming is the proximate location and linkage associated with the humanities. Most of the maps locate the humanities as an appendage to the social sciences or do not locate this field at all. Balaban's (Balaban & Klein, 2006) location of the humanities at the top and the bottom of his hierarchy is quite intriguing, and is supported by anecdote (in his paper), but not by strong analysis. Exactly where should one locate the humanities, and what should the proper linkages be? The data and methodologies used to generate the maps in Table 1 are not well designed to answer these questions. Future maps that better represent the role of the humanities through the use of more comprehensive data sources or better algorithms may shed more light on this question.

DOI: 10.1002/asi

Discussion

The previous section has shown that the consensus map emerges as a noncentric form, and represents the consensus of the 20 maps listed in Table 1 in a suitable fashion (with exceptions noted). In this section, we present the advantages to adopting a noncentric consensus map along with other findings from this study.

Maps of Scholarly Activity

Nineteen of the 20 maps we reviewed are maps of scholarly activity. The maps are intended to describe the relationships between different areas of research. We've also pointed out that these 19 maps do not cover all areas of scholarly work. Early maps of science provided little space for medical research, and one did not include mathematics. The remaining 17 maps are based on the journal literature, and reflect the journal biases in the corresponding databases. For example, we've discussed the fact that the TS database (which underlies most of the maps we've reviewed) does cover the humanities and social sciences, but does not have the necessary depth in more applied areas such as health services and engineering. Efforts to overcome this bias (i.e., using the Scopus database) results in other biases (this database does not cover the humanities).

An unbiased map of scholarly activity can provide insights into regional and national values and beliefs. Scholarly work is not simply about supporting researchers. It's about tradeoffs between supporting the arts, providing an understanding about how people behave, providing health and well-being to society, pursuing technoeconomic goals, and supporting basic research that may have no immediate economic or social impact.

Maps can play an important role in communicating the national orientation towards these different objectives. It is important, however, that the layout does not automatically imply that one area of scholarship is superior to another. Stated differently, one needs a framework where the following seemingly contradictory goals are met: each scholar is at the center of the map, and no scholar is at the center of the map.

The topological solution to this conundrum is to use Riemannian space, to look at circles (where each area is placed on the outside of a circle) or, as in the case of one of the maps we reviewed, to place the areas of scholarship on the surface of a sphere. One of the intriguing aspects of these surfaces is that each node can be thought of as being at the center of the surface (since there are no "ends" to the surface), and at the same time no areas are at the center of the geometry (i.e., the sphere) itself.

This begs the question as to why most maps are generated in Euclidean space. In the previous section on the one-dimensional consensus map, we provided a simple thought experiment that proved that a solution in one-dimensional Riemannian space will often be more accurate than a solution in one-dimensional Euclidean space. This will be true if one increases the dimensionality as well, because the underlying cause (the ability to add in more links) will always appear when one shifts from Euclidean to Riemannian topology.

Despite the inherent advantages of Riemannian topology, there are few if any algorithms or statistical packages that are written for simple forms of Riemannian space. Additionally, we know of no simple test to determine what the underlying Riemannian dimensions of a data set might be, or that would tell us whether the underlying dimensions of a data set are Riemannian or Euclidean. We hold this out as an important need for new algorithms that would be very useful for this area of research.

Maps of Educational Activity

Only one of the 20 maps reviewed here looked at educational activity. Instead of dismissing this map as an outlier, we would like to emphasize that a map based on educational curricula or other indicators of educational activity has some important implications for the way maps of science are created. The following discussion explores some of the implications of creating maps of and/or for educational activity.

Science education is not simply about learning about specific scientific topics or subjects. It should be more about teaching of the scientific method, about the acts of discovery and exploration that enrich our lives in a multitude of areas, about critical thinking, and about the fact that science can and should be fun. Science education should be about communicating that there are many more areas that can be discovered, that students can take part in this process, and that students need not be discouraged with initial difficult experiences with mathematics or physics, or with an ineffective science teacher. (These are not uncommon occurrences among our elementary and secondary students today.)

Maps can play a subtle role in communicating where both old and new discoveries are located. Fundamentally new discoveries are often located in the white spaces, or areas of a map with little or no activity. For example, consider the maps of the world that were drawn in the 13th or 14th century. These maps would show the known world floating on a sea of uncertainty. There would be labels on this open space such as "here there be dragons" to indicate the inherent risks in exploring these areas. The same is true for maps of science. Each map has (or should have) white spaces. These white spaces are where communities of researchers are not located. New discoveries would nearly always appear in the white spaces.

A noncentric map can tell a unique story about the location of new discoveries. In the extreme case where all disciplines are lined up around the circle, new discoveries could be located within the circumference of the circle. Assuming that new discoveries are likely to be more interdisciplinary than past ones, these new discoveries could be located closer to the center of the circle. Their exact locations could reflect the relative influences of existing areas of science on the new work. Placing existing work along the circle, with all points equidistant from a center, implies that all existing work can be

DOI: 10.1002/asi

valued equally. This is not to say that the obviously groundbreaking discoveries of the past have been of no more value to society than other work, but is intended to convey to the student that new discoveries can arise from many directions. A noncentric map can place high value on new discoveries and on interdisciplinary research.

Centric and hierarchical maps, on the other hand, inherently imply that there is more status in some areas of science than in others. A hierarchical map will likely place more emphasis on the role of the top discipline in the hierarchy (usually mathematics). The "hard sciences" have always enjoyed higher status than the so-called "soft sciences." A centric map confers the highest status on the central (dominant) area of research, and also implies that areas of science at the ends of branches are of lower status, or perhaps even dead ends. To us, neither of these two map forms conveys the positive message of science to students in the same way that a noncentric, or status-free, map can.

Use of a noncentric map also follows logically from the differences between classification and mapping that were mentioned earlier. Most classification systems are hierarchical within the various classes. If a higher-level node is assumed, to which all of the classes are linked, these systems then have a centric nature to them. By contrast, algorithms that project multidimensional space down to two dimensions have no centric or hierarchical assumption. Thus, noncentric forms can emerge naturally from mapping efforts, while most classification systems are constrained to hierarchical or centric forms. For these reasons, we would suggest that noncentric consensus maps, such as the ones presented in Figure 5b or 6, would be the most applicable to education.

We also wonder why more maps of science have not been generated from information about educational activity. There is a decided bias towards using the journal literature. While the map generated by Balaban and Klein (2006) may be biased—the data did not reflect what students choose and is greatly influenced by the fact that the college is an agricultural college—the underlying concept is extremely promising. One could generate maps of elementary education, secondary education, and college using different inputs about course curricula and the natural affinities or dislikes that students have for pairs of courses. We hold this out as another important area for future research.

Summary

Our goal in this paper has been to examine many maps of science to see if a consensus map of science would emerge from those data. We have examined, described, and analyzed 20 comprehensive maps of science generated by different researchers, using different data and different methodologies. A consensus map of science did emerge from these data, and takes a noncentric form if all consensus edges, those occurring in over 50% of the input maps, are retained in the map. In addition, the correspondence between the consensus map and of each of the individual input maps was measured. Some shortcomings of the consensus map were also discussed.

We also discussed the inherent advantages of the noncentric form, and showed how use of a Riemannian topology in place of a Euclidean topology would increase the correspondence of a map. In keeping with that finding, we suggest that Riemannian maps of science be used where possible. For instance, the simplest (one-dimensional) solution using a Riemannian topology would be to place the 16 areas of science around the perimeter of a circle. The network of science in this simple case would be represented using 16 nodes and 16 edges. Use of the consensus map in Figure 4b would provide a more complex and more realistic shape, which although not in Riemannian space, is noncentric and thereby has a Riemannian flavor. The correspondence of these simple shapes is extremely good. For instance, the two-dimensional network of Figure 4b captures 75% of all paired relationships from the initial sample of 20 maps. An additional 19% of the paired relationships from the 20 maps are only two hops from each other on the consensus map.

We've pointed out two fruitful areas for future research. First, there is a need for algorithms that are based on Riemannian space. These algorithms have the promise, almost by definition, to generate more accurate maps of science. Second, there is a need to generate more maps of science based on educational activity. Previous work on science mapping is mostly based on the existence of well-structured databases of scholarly activity. Maps based on educational activity can have a significant impact on our understanding of how to teach and communicate the role of science in society. A map of Library of Congress holdings may prove to be very useful, and could have implications far beyond education.

We've also pointed out the difference between science maps and knowledge maps. The maps presented here are not knowledge maps. There is no attempt to answer basic questions of ontology, or to make sure there is a direct and unambiguous correspondence between the phenomena (science) and the classification system. Rather, we address a cartographic question—how one might partition science into a small number of like-sized groups and which groups are adjacent. The fact that these adjacency relationships have not changed dramatically since 1939 is quite remarkable.

We would also like to point out that while this study has produced a consensus map of science, it is not a convergent map of science. There is a substantial difference between the two. The consensus map is the result of an aggregation of information from many different maps. It considers input from experts as well as a variety of algorithmic approaches. Although it is heavily weighted to newer maps, it maintains some contact with maps that are decades old.

By contrast, convergence implies that researchers, over time, are agreeing on partitions (how to divide science), and the explicit linkages between those partitions. A quick look at just three maps published in the last two years, the UCSD map (Klavans et al., 2008), the SCImago-II map (Moya-Anegón et al., 2007), and the co-category map of Leydesdorff and Rafols (2008), shows a lack of convergence. Of these three, one is noncentric, one is centric, and the third is of a mixed form. Due to the variety of data sources, relatedness measures, and mapping algorithms in use today, we do not expect convergence anytime soon.

In lieu of convergence, we suggest that a consensus map can still be very useful for a variety of purposes. We propose that the one-dimensional (Figure 6) and two-dimensional (Figure 4b) consensus maps of science be adopted for educational purposes. The two-dimensional map is more effective for navigational purposes. The one-dimensional map, while it has lower correspondence, is more effective for showing patterns of interdisciplinary research. Given the need to emphasize exploration and discovery, we lean towards adoption of the one-dimensional map, which appears as a simple circle of science.

Acknowledgments

We appreciate the work done by all science mappers, and particularly the work done on those maps included in this study. In particular, we wish to acknowledge that all of us who currently do science mapping are highly indebted to the pioneering work by Henry Small in this area. We also appreciate comments by Peter Hook and several anonymous reviewers that significantly helped to improve this paper.

References

Balaban, A.T., & Klein, D.J. (2006). Is chemistry "The Central Science"? How are different sciences related? Cocitations, reductionism, emergence, and posets. Scientometrics, 69, 615–637.

Bassecoulard, E., & Zitt, M. (1999). Indicators in a research institute: A multilevel classification of journals. Scientometrics, 44, 323–345.

Bernal, J.D. (1939). The social function of science. London: George Routledge & Sons.

Boyack, K.W. (2009). Using detailed maps of science to identify potential collaborations. Scientometrics, 79(1), 27–44.

Boyack, K.W., Börner, K., & Klavans, R. (2009). Mapping the structure and evolution of chemistry research. Scientometrics, 79(1), 45–60.

Boyack, K.W., Klavans, R., & Börner, K. (2005). Mapping the backbone of science. Scientometrics, 64, 351–374.

Comte, A. (1830). Cours de Philosophie Positive (Vol. i).

Ellingham, H.J.T. (1948). Divisions of natural science and technology, The Royal Society Scientific Information Conference (pp. 477–484): The Royal Society, Burlington House. Encyclopaedia Britannica. (2008). Chicago: Author.

Griffith, B.C., Small, H.G., Stonehill, J.A., & Dey, S. (1974). Structure of scientific literatures. 2. Toward a macrostructure and microstructure for science. Science Studies, 4, 339–365.

Klavans, R., & Boyack, K.W. (2006a). Identifying a better measure of relatedness for mapping science. Journal of the American Society for Information Science and Technology, 57, 251–263.

Klavans, R., & Boyack, K.W. (2006b). Quantitative evaluation of large maps of science. Scientometrics, 68, 475–499.

Klavans, R., & Boyack, K.W. (2007). Is there a convergent structure of science? A comparison of maps using the ISI and Scopus databases. In D. Torres-Salinas & H. Moed (Eds.), Proceedings of the 11th International Conference of the International Society for Scientometrics and Informetrics (pp. 437–448). Madrid, Spain: ISSI.

Klavans, R., & Boyack, K.W. (2008). Thought leadership: A new indicator for national and institutional comparison. Scientometrics, 75, 239–250.

Klavans, R., Boyack, K.W., & Patek, M. (2008). Forecasting large trends in science using a multi-database map of journals. Unpublished manuscript.

Leydesdorff, L., & Rafols, I. (2008). A global map of science based on the ISI subject categories. Journal of the American Society for Information Science and Technology (in press, DOI:10.1002/asi20967).

Martin, S., Brown, W.M., Klavans, R., & Boyack, K.W. (2008). DrL: Distributed recursive (graph) layout. Sandia National Laboratories report, SAND 2008-2936J

Marris, E. (2006). 2006 Gallery: Brilliant display. Nature, 444, 985–991.

Moya-Anegón, F., Vargas-Quesada, B., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., Munoz-Fernández, F.J., & Herrero-Solana, V. (2007). Visualizing the marrow of science. Journal of the American Society for Information Science and Technology, 58, 2167–2179.

Moya-Anegón, F., Vargas-Quesada, B., Herrero-Solana, V., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., & Munoz-Fernández, F.J. (2004). A new technique for building maps of large scientific domains based on the cocitation of classes and categories. Scientometrics, 61, 129–145.

Narin, F., Carpenter, M., & Berlt, N.C. (1972). Interrelationships of scientific journals. Journal of the American Society for Information Science, 23, 323–331.

Rosvall, M., & Bergstrom, C.T. (2008). Maps of random walks on complex networks reveal community structure. Proceedings of the National Academy of Sciences of the USA, 105, 1118–1123.

Samoylenko, I., Chao, T.-C., Liu, W.-C., & Chen, C.-M. (2006). Visualizing the scientific world and its evolution. Journal of the American Society for Information Science and Technology, 57, 1461–1469.

Shiffrin, R.M., & Börner, K. (2004). Mapping knowledge domains. Proceedings of the National Academy of Sciences of the USA, 101(Suppl 1), 5183–5185.

Small, H. (1999). Visualizing science by citation mapping. Journal of the American Society for Information Science, 50, 799–813.

Small, H., & Garfield, E. (1985). The geography of science: Disciplinary and national mappings. Journal of Information Science, 11, 147–159.

Spencer, H. (1864). Illustrations of universal progress: A series of discussions. New York: D. Appleton & Company.

Appendix A: Matrix of Edges by Map

Edge	Bernal	Ellingham	Balaban-I	Small74	Small85	Small99	KB-Para	KB06-TS	KB06-SC	B-Z	K02	Backbone	BBK02-S	B03-ST	UCSD	SCImago-I	SCImago-II	L-R	Rosvall	Balaban-II	Total
B;I-BC						1															1
B;I-C						1															1
B;I-MD;N						1															1
B;I-PS						1															1
B-B;I						1															1
B-BC	1	1			1		1	1	1	1	1	1	1	1	1	1		1			14
B-BC;I			1																1	1	3
B-BC;I;MD				1													1				2
B-C	1		1			1					1										4
BC;I;MD-C																	1				1
BC;I;MD-CS				1																	1
BC;I;MD-I				1													1				2
BC;I;MD-M				1																	1
BC;I;MD-MD				1													1				2
BC;I;MD-N				1													1				2
BC;I;MD-P				1																	1
BC;I;MD-PC				1																	1
BC;I;MD-PS				1																	1
BC;I;MD-SS				1																	1
BC;I-C;PC																			1		1
BC;I-E																			1		1
BC;I-MD																			1		1
BC;I-N																			1		1
BC;I-PS;SS			1																		1
B-C;PC				1																1	2
BC-BC;I;MD				1													1				2
BC-C	1	1				1	1	1	1			1		1	1						9
BC-C;P;PC						1															1
BC-C;PC				1	1											1					3
BC-CS				1														1			2
BC-G						1															1
BC-HS;I;MD																1					1
BC-HS;MD		1																			1
BC-I				1			1	1	1		1	1		1	1			1			9
BC-I;MD;N	1																				1
BC-MD					1		1	1	1	1		1			1	1		1			9
BC-N											1	1						1			3
BC-N;PS										1											1
BC-PC								1													1
B-E	1	1								1					1				1		5
B-G	1	1	1		1		1	1	1	1	1	1	1	1	1		1	1	1	1	17
B-HS;I;MD																1					1
B-I		1					1	1		1			1								5
B-I;MD;N	1																				1
B-M																	1				1
B-MD					1				1												2
B-N			1																	1	2
B-P											1										1
B-PS;SS					1																1
C;P;PC-E						1															1
C;PC-CS				1																1	2
C;PC-E					1											1			1		3
C;PC-G																1				1	2
C;PC-P				1	1											1			1	1	5
C;PC-PC				1																	1

(*Continued*)

DOI: 10.1002/asi

Appendix A. (*Continued*)

Edge	Bernal	Ellingham	Balaban-I	Small74	Small85	Small99	KB-Para	KB06-TS	KB06-SC	B-Z	K02	Backbone	BBK02-S	B03-ST	UCSD	SCImago-I	SCImago-II	L-R	Rosvall	Balaban-II	Total
C-E	1	1					1		1						1		1				6
C-G	1		1			1	1	1	1			1					1				8
C-HS;MD		1																			1
C-MD										1											1
C-N;PS										1											1
C-P	1		1										1	1				1			5
C-PC		1	1				1	1	1	1	1	1	1	1	1		1	1			13
CS-E	1	1	1				1		1				1			1		1		1	9
CS-G	1																				1
CS-H;SS							1	1			1										3
CS-M		1					1	1	1	1	1	1	1	1	1	1		1		1	13
CS-MD																			1		1
CS-N							1														1
CS-P	1	1			1								1	1	1	1			1		8
CS-PC	1		1														1				3
CS-SS	1								1			1	1	1	1				1		7
E-G	1	1	1		1	1	1	1	1			1		1	1	1	1	1	1	1	16
E-M		1					1	1	1	1			1	1				1			8
E-MD																			1		1
E-P		1						1	1	1		1	1	1	1	1	1		1		11
E-PC		1					1		1	1		1	1	1	1		1	1			10
E-PS;SS			1																		1
E-SS	1																1				2
G-M				1				1													2
G-P					1		1			1	1		1						1		6
G-PC	1																				1
G-PS;SS			1																		1
H;SS-N											1										1
H;SS-PS							1	1			1										3
H-M			1																		1
H-PS;SS	1		1																		2
HS;I;MD-MD																1					1
HS;I;MD-N			1																	1	2
HS;I;MD-PS;SS			1																	1	2
HS;MD-I		1																			1
HS-I					1																1
HS-I;MD;N	1																				1
HS-MD					1		1	1	1	1	1		1	1	1		1	1			11
HS-N								1	1			1	1	1							5
HS-PS							1					1	1		1						4
HS-PS;SS	1				1																2
H-SS															1	1	1				3
I;MD;N-PS;SS	1																				1
I-MD					1		1	1	1	1	1	1	1	1	1			1			11
MD;N-PS						1															1
MD-N					1		1	1	1		1	1			1			1	1		9
MD-N;PS										1											1
MD-PS																1		1	1		3
M-MD;N						1															1
M-P		1	1		1		1	1	1		1	1			1				1	1	11
M-PS						1															1
M-PS;SS																				1	1
M-SS				1										1		1	1				4
N-PS				1			1	1	1		1	1	1	1	1		1		1		11
P-PC	1	1		1			1	1	1		1	1	1	1	1		1	1			13
PS;SS-SS	1																				1
PS-SS						1			1			1	1	1	1	1	1		1		9

DOI: 10.1002/asi