Jointly published by Akadémiai Kiadó, Budapest
and Springer, Dordrecht

Scientometrics,
Vol. 66, No. 2 (2006) 377–387

Combining concept maps and bibliometric maps: First explorations

R. K. BUTER,[a] E. C. M. NOYONS[a] M. VAN MACKELENBERGH,[b] T. LAINE[c]

[a] *Centre for Science and Technology Studies (CWTS), Leiden University, Leiden (The Netherlands)*
[b] *Rosmalen (The Netherlands)*
[c] *University of Joensuu, Joensuu (Finland)*

Bibliometric maps of science are a well-established research subject. But their adoption as a science policy support tool is lacking. We think this is because the user does not immediately comprehend a map and (as a result) is not enticed into using it. To help this comprehension, we propose the use of "qualitative maps": an umbrella term for diverse tools such as concept maps and mental maps. We developed a tool that interfaces between a qualitative map and a bibliometric map which lets the user create a correspondence between the distinct vocabularies of the maps. We also conducted two user studies: the first explored the combined use of bibliometric and qualitative maps and the second the preferred format of the map and the word-usage in the description of its elements.

Introduction

Maps of science have been around for quite a while. They are firmly established as a bibliometric and scientometric research subject, as illustrated by works like CALLON et al. (1986), BRAAM et al. (1991), WHITE & MCCAIN (1998) and NOYONS (1999). Mapping can be considered a bridge to related research in fields such as Knowledge Visualization (e.g. CHEN, 1999) and Information Retrieval (e.g. LIN, 1998).

Bibliometric maps are graphical summaries of sets of papers, either based on citation data, words or phrases or some other bibliometric elements. Part of their appeal is based on the shared characteristics with their geographic counterparts (SMALL, 1999) allowing them to be referred to as "landscapes of science" (NOYONS, 1999).

Besides the maps deployed in Bibliometrics, a number of (diagrammatic) representation schemes has been developed with a distinctive character. These schemes have names such as *mental maps*, *mind maps* or *concept maps*. The appearance of these maps varies: some are like networks, either with directed or undirected relations; some have labelled relations; some are strictly hierarchical; and some are completely free-form. The position of the elements and their relations are laid out manually and are

Received June 22, 2005

Address for correspondence:
R. K. BUTER
Centre for Science and Technology Studies (CWTS), Leiden University, Leiden, The Netherlands
E-mail: buter@cwts.leidenuniv.nl

0138–9130/US $ 20.00

usually from the hand of a single person. In contrast, the structure of bibliometric map is generated algorithmically and supposedly without user bias. We will use the term *qualitative map* (QM) to refer to these different, but related mapping techniques.

Qualitative maps can, for example, be used to take notes in a lecture or a meeting or lay out knowledge to benefit education (NOVAK, 1998). And when carefully crafted to represent a cognitive structure of some problem matter, the depictions can also be analysed and used as an inference tool. Examples of matters being analysed this way include: *strategies* (HUFF, 1990) *organizational learning* (SRINIVAS & SHEKAR, 1997) and *international politics* (YOUNG, 1996).

Even though the research in bibliometric maps has settled itself nicely, the adoption of maps as a policy-supporting tool is lacking. Early problems described by HEALEY et al. (1986) are partly to blame. But techniques have advanced and problems have been addressed. Yet even recent efforts such as the EU-funded Centres of Excellence project[1] that, among other things, explicitly targeted the further development of bibliometric maps as a science policy support tool have not been able to change this situation.

We have extensive experience with maps in combination with evaluative bibliometric analyses, as illustrated in e.g. NOYONS et al. (1999) and NOYONS (2004). Based on this experience, we can say there is not much wrong with the *potential* of maps. After having been guided through the map, a knowledgeable user is usually quick to see advantages of the use of a bibliometric map (BM). Yet without such guidance, there is less response from users. We could anthropomorphically speak of a “lack of communication” between a BM and its users. And this *communication* problem is an important reason for the lacking adoption of bibliometric maps as a science policy tool.

As LAW (1986) suggests, an author writing a paper should carefully phrase that paper in order to get and *keep* the attention of a potential reader. The phrasing should be broad enough to attract as many interested readers as possible, yet specific enough to avoid the informed reader to dismiss it as being without value. Maps of science are like scientific papers in this respect: a user should be attracted by the broad overview given by the map and find valuable information in its details.

Since most of our maps are built with textual elements, i.e. with words and phrases taken from papers, we could be using the wrong words to describe (label) the elements in our maps. Then, much like an author rephrasing parts of a paper to increase its readability, we could change the descriptions of particular elements to improve the “communication” between a user and a map. But maps don’t have the same verbosity as papers do: only short phrases fit on the limited amount of space. And variance in vocabulary between users is too large, as described by FURNAS (1987), to find a set of labels that fits all. As a result, it will be next to impossible to find a labelling that will be descriptive to most (let alone *all*) users.

[1] For more information, we refer to the website ‘studies.cwts.nl/projects/ec-coe’.

We expect a QM can help here and especially the *process* of creating one: the process *activates* the user's vocabulary while the user searches for words to describe the elements, at the same time *materializing* a user-specific vocabulary in these descriptions. The result is a personal, qualitative description of a problem matter: a depiction of elements relevant to the *user*, bearing labels that are meaningful to that *user* and the relations between the elements the *user* considers important. If we can then map the terms in this description to terms on the bibliometric map, we have created links between the (materialized) internal representation of the user and the bibliometric map.

The research described in this paper explores this idea. But since the research is still in progress, it leaves many technical and fundamental questions untouched. Two main components of our explorations are presented. The first is a tool that acts as an intermediary between a QM and a BM; this tool lets a user create trees of associated terms, with at the root of those trees terms extracted from the qualitative map. The second component is a set of two distinct user studies: the first study investigated how a BM is used in combination with a QM and the second study investigated how the preferred QM of a user may look like.

Methods and tools

Woven Stories

To give our research a head start, we used an existing implementation of a qualitative mapping tool. This tool was (and still is) developed to explore the utility and applicability of "Woven Stories". The concept of Woven Stories (WS) was originally created to support collaborative writing over the web. But more generally, it can be regarded as a cognitive support tool (GERDT, 2001). It is applicable to different domains, like for example E-Learning or Corporate Strategy planning (NUUTINEN, 2004).

With WS it is possible to write documents with different authors in a distributed, collaborative manner. This is different from, for example, circulating different versions of the same document between different authors, where one has the problem that only a single author can, at any point in time, be working on the "active" version of the document. WS tries to remedy this by dividing the document into sections to which a user may add comments or which a user can edit. These sections are all stored on a central server, which distributes changes to all participating in ("logged on" to) the story. If the user edits a section, a new version is created and communicated to all other users, allowing different versions of the same section to be written parallel.

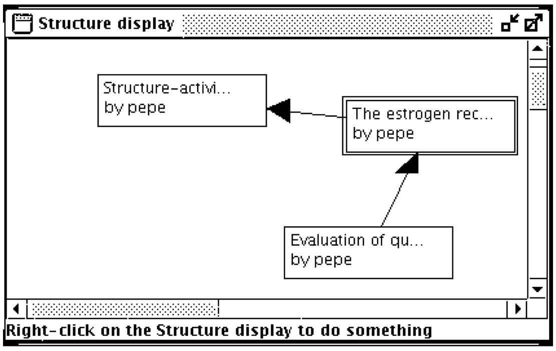

Figure 1. A window from a Woven Stories application, showing three sections (each having a single version) and a single path between these sections. All sections have the fictitious author "pepe".

The different section versions are represented as nodes in a directed graph, as illustrated in Figure 1. By following a certain path through the section-graph, a specific version of the document (or "story") is created. The final document is a specific section path.

We used the sections of WS as the elements of a qualitative map and the paths between those sections as relations between the elements. And, as can be seen from Figure 1, the relations are directed and unlabelled. Additional features of the WS tool, such as the distributed nature and the ability to add comments to sections, were not used.

Bibliometric mapping application

The bibliometric mapping application "*i*BeX" (Interactive Bibliometric EXplorer) we presented in BUTER & NOYONS (2001) is a mix between an information retrieval and a visualization tool. For example, it can retrieve the list of authors with the highest impact in a certain element or group of elements on a map. But it can also visualize that list by colouring map elements, or display different types of charts either on the elements or on the map as a whole. This application is actively deployed in evaluative bibliometric studies, like the already mentioned Centres of Excellence project. The application had to be slightly extended to be able to communicate with the tool presented below.

Correspondence interface

The choice of words in the descriptions of the QM (its *vocabulary*) is not necessarily the same as that in the BM. Again, this is because there are many ways a thing or idea can be described. So, multiple related terms are needed to properly describe the research and to retrieve all relevant papers from the database. So, what is needed is a way to map the words or phrases used in a QM to the corresponding ones in the BM.

This correspondence-creating interface, or simply correspondence interface (CI), *interactively* creates relations between words or phrases from the QM and those of the BM. First, the CI extracts words and phrases from the QM: these are the candidate terms. The user can select some of these terms into a working set, with of course the possibility to extend this working set later on. The correspondence is developed by expanding the terms in the working set into associated terms from the database of the BM. These expansion terms can also be expanded, resulting in an expansion *tree*.

This is actually something akin to what in Information Retrieval is known as Query Expansion (e.g. XU & CROFT, 1996), where more and more specific queries are created, each one getting closer to retrieve some desired result. However, instead of a query to retrieve documents, a *correspondence* between the vocabulary used in the QM and the one used in the BM is created. And the result is not a (set of) documents, but a proper coverage of papers (and hence, themes or elements) in the BM.

First user study

The objective of the first user study was to investigate how a QM changes the perception and use of a BM. This was done by employing *user scenarios*. User scenarios describe the *user*, the *goal* of the user, the *task* the user engages in to reach the goal, the *conditions* under which the user performs the task and the *criteria* that determine whether the result (of the task) can be considered a success. All this is formally described in TC 159/SC 4 (2000).

The user scenario we designed for this study, targeted field experts, such as researchers and people involved in science policy. The goal they were given was to search and compare actors considered relevant to a specific research area. To accomplish this, the users had the task to use a BM and to describe the research in a QM, which had to be drawn using some generally available, third-party piece of QM software. For this, the description of the research in the QM must be (considered) in agreement with the research as displayed in the BM. In a successful task, the user confirms the actors are located in the expected parts on the BM.

Second user study

The objective of this study was to investigate the preferred format of a QM and the relation between the word-usage in the BM and the QM. The set-up of this study was rather low-tech: an A3-sized sketchbook and a permanent marker. We did not use a user scenario for this study; instead, we simply instructed the participants, all of them working in the field of Bibliometrics and Scientometrics, to draw a depiction of "bibliometric analysis" on a piece of A3 paper using that permanent marker. To illustrate what a possible result could look like, the participants were shown a depiction of a network of terms. At the same time, the participants were specifically instructed to format their depiction as *they* saw fit, possibly disregarding the example altogether.

Each participant was given ten minutes to complete the depiction. Somewhere in the last five minutes, the participants were asked to point out their research in the description and to improve the depiction to more accurately represent this research. The two main reasons for these additional questions were to avoid very general descriptions and to direct users who were not sure how to continue.

Implementation and results

Correspondence Interface

The implementation of a prototype CI resulted in the tool depicted in Figure 2. Shown is a working set of five root terms, four of which have been expanded. For example, the tree at the top of the window, sprouting from the root term "xenochemical" has been expanded into three associated terms: "hormone", "milieu" and "contaminant". The last expansion term has also been expanded into two associated ones: "concentration" and "chemical".

Elements in the CI can be collapsed into a single node, hiding the subtree of that node. To avoid confusion between unexpanded and collapsed elements, an element is given a "+" to indicate it has been expanded and possibly only partly shown.

Expansion terms are selected from a dialog that contains a list of candidate terms, sorted to some criterion. The algorithm that sorts these lists of candidate expansion terms was very straightforward in this prototype: it returned the terms in papers containing the term that is expanded, sorted on frequency of occurrence.

To test the relation between a correspondence and a BM, we implemented two views on the BM:

1. a pie-chart per map-element to show the coverage per selected term in that element;
2. a spread of colours over the map to show the coverage of the complete selection.

Figure 3 shows an example of the first type of view: pie-charts on a map of the field "Endocrine Disruptors" for four selected terms. Although "chemical", "endocrine" and "concentration" are spread in relatively equal proportions over the map, the occurrence of "contaminant" seems to be less frequently used in papers in the south-east of the map. However, a clear picture does not emerge.

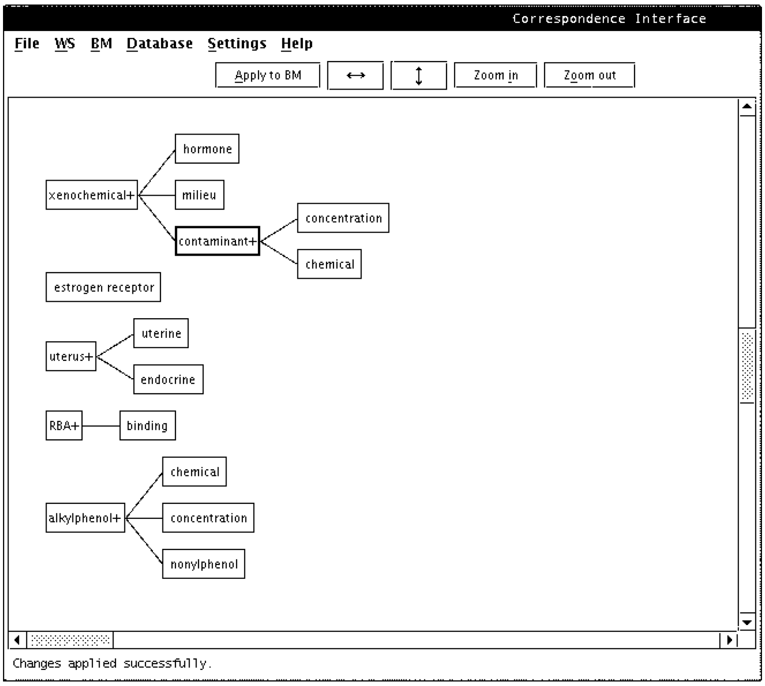

Figure 2. The CI with five root terms, four of which are expanded. The expansion term "contaminant" has also been expanded.

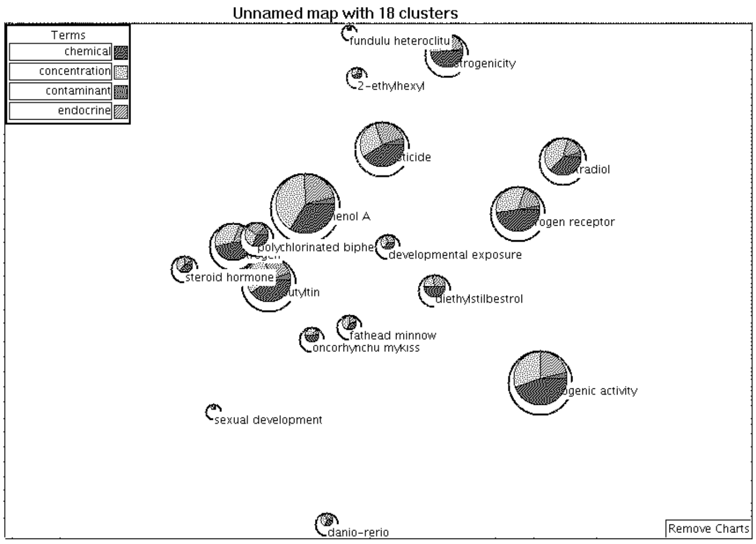

Figure 3. A co-word map in *i*BeX of the field "Endocrine Disruptors" decorated with pie-chart showing the per-element proportional number of papers of four expansion terms.

First user study

Unfortunately, it turned out to be hard to find suitable and *available* experts: only one participant was found, a female researcher in the field of "Endocrine Disruptors". What follows is a description of the execution of the scenario for this participant.

We presented the participant a BM of her field and asked her to assess it. And although there was criticism, the vast majority of the clusters on the map were deemed meaningful in her field.

Then, we asked the participant to create a map that described her field. She divided the research into a limited number of main subjects, which she distributed more or less evenly over the map. To these main subjects she then added associated, related descriptions and drew the relevant inter-subject relations. Notable was the increased focus on the themes relevant to the research of the participant herself, expressed by adding more details to that part of her map.

After completing the QM, the participant was asked to look at the BM again and describe how it related to her map. All the main subjects from the QM were represented in the map. However, now she found the BM lacking: it was either too detailed or too coarse. As a result, she judged the map not trustworthy enough to continue with the rest of the task.

Second user study

Although there was no generally preferred format, six out of eight participants based their map on a network of descriptions. The other two based it on either a table or a hierarchy. Most participants also *grouped* descriptions into super-structures, which could have relations to other elements, indicating such relations were applicable to all elements of the group.

The descriptions used in the maps were collected for further analysis. Most descriptions consisted of only one word or phrase, but more complex descriptions were cut into a list of words and phrases (or *terms*, below). Some terms had to be translated from Dutch to English; the terms that turned out to be untranslatable were discarded.

After translation and spelling unifications (like for example *visualisation*→*visualization*) there were 121 distinct terms. The number of distinct terms per user was between 13 and 29, with a median of 19. Of the 121 terms, only 20 (16%) were used more than once. The most frequent term was "bibliometric analysis" (the term given to the participants to describe), appearing in 5 depictions. The database contained 81 out of the 121 terms. For each participant, between 52 and 93 percent of the terms could be found in the database, with a median of 74%. Of the terms that could not be found, the majority were not nouns or noun phrases, but for example adjectives, which are not kept readily retrievable in the BM database.

There were 132 different words used in the descriptions. The database contained 122 (92%) of these words. There were 33 words which appeared more than once. Per participant, there were between 19 and 36 distinct words, with a median of 21.50. Per phrase, there were between 0.65 and 1.0 unique words, with a median of 0.78.

We also did a quick analysis of the relations used in the maps. We used only the relations explicitly drawn by the user, not the ones implied by the vicinity of descriptions. A relation was represented by a pair of terms. If a relation was drawn between groups of descriptions, or if a complex description was converted to a list of terms, all possible pairs were used. For example, a relation between a single term and a group of three terms resulted in three distinct pairs. Of course, this inflated the number of pairs considerably, possibly over-representing some terms. For example, in the set of pairs, the most frequently used term was "citation", appearing in ten pairs, but used by only three participants. The result was a list of 121 distinct pairs. The number of pairs per user was between 9 and 25, with a median of 13.5. In these pairs, 90 distinct words were used, 58 appearing more than once, and ten appearing more than five times.

Discussion and conclusions

We presented our first explorations in the use of concept mapping and related techniques to improve the comprehension and usage of bibliometric maps (BMs). This research is part of our on-going investigation into improving the use of bibliometric maps as a science policy tool. We used the term Qualitative Map (QM) as an umbrella term to refer to the different techniques related to concept mapping.

A supportive observation from the first user study is the increased level of scrutiny *after* having drawn a qualitative map of the field. This seems to corroborate with the idea that the creative process producing a QM indeed increases focus which allows a more scrutinized judgment of a BM and possibly improving its comprehension. In this study however, this increased scrutiny also invalidated the criterion set for a successful scenario: the user did not trust the BM provided a valid structure of the research.

A probable reason for this failure is the naive field delineation for "Endocrine Disruptors", which essentially boiled down to papers matching the search "endocr* disr*". Unfortunately, there was no opportunity to improve this delineation and repeat the exercise. But a proper and complete field definition will always be important. The second user-study showed that a majority of the descriptions from a QM can be found in the database of the BM, but this most likely depends on that field delineation: if complete, the user-vocabulary will be available in the papers they are familiar with or even contributed to. However, if that field-delineation is lacking, terms preferred by the user may very well be missing.

The network format seems to be the most preferred format to use in a BM. Of course, it cannot be ruled out that the example given to the participants, which showed a network of terms, introduced a bias towards the network format. Also interesting is the use of groups (of terms and relations) in the depictions, something not commonly seen in the QM related literature and existing QM tools.

Creating a correspondence between the vocabulary of the QM and that of the BM is required for a proper coverage of the research. The representation of this coverage as a forest of trees has an obvious advantage in allowing multiple senses to be present in different trees in the same correspondence. However, in a specialized research area this is probably rarely needed. And also, if a root term is not found in the database, there is in the current implementation no way of expanding that term. If we changed the representation into a network, a user might be able to relate it to terms already expanded from other terms. Such related terms could than be used as clues for approximate matching.

By letting a user create a QM of research relevant to a given science field, we think the comprehension of a BM can improve and as a result make the BM a more attractive tool. But the kind of problem depicted by the user is not necessarily restricted to the delineation of research. Therefore, a possible future extension could be to ask someone to draw up a description of an interesting current research theme, or of actors and their research focus as perceived by that person and *directly* decorate such depictions with bibliometric indicators.

Still, for the moment there is still a long way to go before this combined mapping approach is usable for inclusion in the policy supporting tool chain. Nevertheless, we are convinced the ability to materialize qualitative expert knowledge and use it in a bibliometric analysis, even beyond being an assistance in comprehending a BM, has a lot of potential and deserves further attention.

*

We would like to thank An Ghekiere for participating in the first user study. The research described in this paper was conducted as part of the EC-funded research project called "Monitoring Environment and Health R&D in Relation to Socio-Economic Problems: a New Approach to Impact Assessment" (contract number QLK4-CT-2002-30244).

References

BUTER, R. K., NOYONS, E. C. M. (2001), Improving the functionality of interactive science maps. *Scientometrics*, 51 : 55–68.

CALLON, M., LAW, J., RIP, A., (Eds) (1986), *Mapping the Dynamics of Science and Technology, Sociology of Science in the Real World*, The Macmillan Press Ltd., London.

CHEN, C. (1999), Visualising semantic spaces and author co-citation networks in digital libraries, *Information Processing and Management*, 35 : 401–420.

FURNAS, G. W., LANDAUER, T. K., GOMEZ, L. M., DUMAIS, S. T. (1987), The vocabulary problem in human system communication: An analysis and a solution, *Communications of the ACM*, 30 (11) : 964–971.

GERDT, P., KOMMERS, P., LOOI, C.-K., SUTINEN, E. (2001), Woven Stories as a cognitive tool. In: BEYNON, M., NEHANIV, C. L., DAUTENHAHN, K. (Eds), *Cognitive Technology 2001*, *Lecture Notes in Artificial Intelligence 2117,* Springer-Verlag, Berlin, pp. 233–247.

HUFF, A. S. (Ed.) (1990), *Mapping Strategic Thought*, John Wiley & Sons Ltd., Chichester.

LAW, J. (1986), The heterogeneity of texts. In: CALLON (1986), pp. 67–83.

LIN, X. (1998), Map displays for information retrieval, *Journal of the American Society for Information Science*, 48 (1) : 40–54.

NOVAK, J. D. (1998), *Learning, Creating and Using Knowledge, Concept Maps as Facilitative Tools in Schools and Corporations*, Lawrence Erlbaum Associates, New Jersey.

NOYONS, E. C. M., LUWEL, M., MOED, H. F. (1999), Combining mapping and citation analysis for evaluative bibliometric purposes, *Journal of the American Society for Information Science*, 50 : 115–131.

NOYONS, E. C. M. (1999), *Bibliometric Mapping as a Science Policy and Research Management Tool*, PhD Thesis, Centre for Science and Technology Studies, Leiden University, DSWO Press, Leiden.

NOYONS, E. C. M. (2004), Science maps within a science policy context. In: MOED, H., GLÄNZEL, W., SCHMOCH, U. (Eds), *Handbook of Quantitative Science and Technology Research, The use of Publication and Patent Statistics in Studies of S&T Systems*, Kluwer Academic Publishers, Dordrecht, pp. 237–255.

NUUTINEN, J., SUTINEN, E., LIINAMAA, K., VANHARANTA, H. (2004), Woven Stories as a tool for corporate strategy planning. In: *Proceedings of the WBC2004*, IADIS Press, Lisbon, pp. 438–441.

SMALL, H. (1999), Visualizing science by citation mapping, *Journal of the American Society for Information Science*, 50 (9) : 799–813.

SRINIVAS, V., SHEKAR, B. (1997), Applications of uncertainty-based mental models in organizational learning: A case study in the Indian automobile industry, *Accounting, Management and Information Technologies*, 7 (2) : 87–112.

TC 159/SC 4 (2000), *Human-Centred Lifecycle Process Descriptions*, ISO standard 18529.

WHITE, H., MCCAIN, K. (1998), Visualizing a discipline: An author co-citation analysis of information science, 1972-1995, *Journal of the American Society for Information Science*, 49 (4) : 327–355.

XU, J., CROFT, W. B. (1996), Query expansion using local and global document analysis, *Proceedings of the Nineteenth International ACM SIGIR Conference*, pp. 4–11.

YOUNG, M. D. (1996), Cognitive mapping meets semantic networks. *Journal of Conflict Resolution*, 40 (3) : 395–414.