

Bibliometric analysis of two subdomains in philosophy: free will and sorites

Per Ahlgren · Peter Pagin · Olle Persson · Maria Svedberg

Received: 10 June 2014 / Published online: 15 February 2015
© Akadémiai Kiadó, Budapest, Hungary 2015

Abstract In this study we tested the fruitfulness of advanced bibliometric methods for mapping subdomains in philosophy. The development of the number of publications on free will and sorites, the two subdomains treated in the study, over time was studied. We applied the cocitation approach to map the most cited publications, authors and journals, and we mapped frequently occurring terms, using a term co-occurrence approach. Both subdomains show a strong increase of publications in Web of Science. When we decomposed the publications by faculty, we could see an increase of free will publications also in social sciences, medicine and natural sciences. The multidisciplinary character of free will research was reflected in the cocitation analysis and in the term co-occurrence analysis: we found clusters/groups of cocited publications, authors and journals, and of co-occurring terms, representing philosophy as well as non-philosophical fields, such as neuroscience and physics. The corresponding analyses of sorites publications displayed a structure consisting of research themes rather than fields. All in all, both philosophers involved in this study acknowledge the validity of the various networks presented. Bibliometric mapping appears to provide an interesting tool for describing the cognitive orientation of a research field, not only in the natural and life sciences but also in philosophy, which this study shows.

P. Ahlgren (✉)
School of Education and Communication in Engineering Sciences (ECE), KTH Royal Institute
of Technology, 100 44 Stockholm, Sweden
e-mail: perahl@kth.se

P. Pagin · M. Svedberg
Department of Philosophy, Stockholm University, 106 91 Stockholm, Sweden
e-mail: peter.pagin@philosophy.su.se

M. Svedberg
e-mail: maria.svedberg@philosophy.su.se

O. Persson
Department of Sociology, Inforsk, Umeå University, 901 87 Umeå, Sweden
e-mail: olle.persson@soc.umu.se

Keywords Bibliometrics · Cocitation analysis · Free will · Mapping · Philosophy · Sorites

Introduction

An important theme within bibliometric research is science mapping, where maps in the form of networks are constructed, with the aid of suitable programs (e.g., Leydesdorff and Persson 2010; Leydesdorff and Rafols 2011; Meyer et al. 2014; Shen et al. 2013; Tatry et al. 2014). The purpose of a mapping study is often to uncover the cognitive structure of a given research field, and the results of such studies are potentially useful for policy makers. In this study, which falls under the theme science mapping, we focus on a field within the humanities, namely philosophy.

We have found only a few bibliometric studies that treat philosophy. Citations from philosophy dissertations to journals and serials were analyzed by Herubel (1991) in the context of library collection development. Cullars (1998) studied citations occurring in English-language monographs in philosophy, and found that the citation patterns were typically humanistic, with the main part of the citations to books rather than to journal articles. Hyland (1999) analyzed, among other things, the form of citations occurring in research articles in eight fields, among them philosophy. It was found, for instance, that all fields but philosophy displayed a preference for non-integral citations, i.e., citations that make reference to the author in parenthesis or by superscript numbers.

At the level of research specialties, Kreuzman (2001) made a study of particular relevance for our paper. A (first author) cocitation analysis was performed of 62 preselected authors in philosophy with the objective to throw light upon the relation between philosophy of science and epistemology. Based on normalized cocitation data, multidimensional scaling and cluster analysis were used to map the authors. Lindholm-Romantschuk and Warner (1996) used citation analysis to examine the diffusion of ideas over time, particularly as these are embodied in monographs, in philosophy, sociology and economics. The findings indicated, for instance, that monographs generate a substantially higher number of citations than journal articles in the three fields studied. Citations from journal articles in eight humanistic fields, among them philosophy, were analyzed by Knievel and Kellsey (2005), who found that, in most of the studied fields, citations to books dominated the citation links. It was also found, however, that citation patterns varied widely among the eight fields. Publication output and received citations for 12 internationally renowned researchers in five fields, including philosophy, were analyzed by Baneyx (2008). This author used two databases, Google Scholar and Web of Science, and analyzed result differences between them. Kabelka (2012) analyzed growth of Lithuanian philosophical discourse and observed an exponential growth rate since 1970, whereas Manana-Rodriguez and Gimenez-Toledo (2013) detected, using a quantitative approach, significant differences between Spanish psychology and philosophy journals.

A number of bibliometric works have analyzed various aspects of humanistic research, without special reference to philosophy. Thompson (2002) studied citation patterns in literary scholarship, whereas Larivière et al. (2006) compared natural sciences and engineering with social sciences and humanities with regard to citation practices. Hammarfelt (2011a) studied the impact and the dissemination of a highly cited publication, Walter Benjamin's *Illuminations*, as well as citation patterns in Swedish literary studies

(Hammarfelt 2012). One of the few, and first, studies looking at the possibility of mapping topics within the humanities was made by Leydesdorff and Salah (2010). These authors presented several citation networks with journals as nodes. It was shown that art journals were cited by (social) science journals more than by other art journals, but also that these art journals draw upon one another in terms of their own references. Interdisciplinary aspects of humanistic research have been studied also by other authors (Dowell 1999; Hammarfelt 2011b), and the journal structures in the Arts & Humanities Citation Index (A&HCI), one of the citation databases in Web of Science, have been mapped (Leydesdorff et al. 2011).

The purpose of this study is to test the fruitfulness of advanced bibliometric methods for mapping subdomains in philosophy. We treat the two subdomains *free will* (FW) and *sorites* (SOR),¹ and we put forward the following research questions:

1. How has the number of publications on FW (SOR) developed over time?
2. Which are the most frequently cited *publications* in publications on FW (SOR), and how are these cited publications related in terms of cocitations?
3. Which are the most frequently cited *authors* in publications on FW (SOR), and how are these authors related in terms of cocitations?
4. Which are the most frequently cited *journals* in publications on FW (SOR), and how are these journals related in terms of cocitations?
5. What can we say if we map terms that occur frequently in publications on FW (SOR), using a term co-occurrence approach?

We will try to shed light on these questions by using publication data from Web of Science, provided by Thomson Reuters. This database indexes foremost journals. However, philosophy—compared to many other fields in the humanities—has a strong tradition of publishing in journals and is therefore better suited for using journal oriented databases.

The remainder of this paper is organized as follows. Short introductions to the two subdomains, FW and SOR, are given in the next section. Data and methods are treated in the third section, while the following section reports the results of the study. The final section contains a discussion, as well as conclusions.

FW and SOR: short introductions

FW

The question of whether we have free will is one of the most widely discussed in the modern philosophical debate and there are many problems that are relevant to this question. Some issues about free will were raised and discussed already by ancient philosophers like Aristotle, the Epicureans and the Stoics.

One important issue is whether free will is compatible with determinism or not. Determinism is the thesis that everything that happens is completely determined by events in remote past together with the laws of nature. Contrary to popular belief, the falsity of

¹ The two bibliometricians of the study did some preliminary analyses of FW publications, and they decided to ask a philosopher with expertise in FW to participate in the study as a domain expert. For SOR, one of the bibliometricians informed a university colleague and philosopher about the study, and this philosopher was asked to select a theme within the philosophy of language, the main interest of the philosopher, and to participate in the study as a domain expert.

determinism has not been proven by quantum mechanics and the truth is that we do not know today whether or not our world is deterministic.² The position that free will is incompatible with determinism is called *incompatibilism* and the opposite view is labeled *compatibilism*. One of the most influential incompatibilist arguments is Peter van Inwagen's Consequence Argument which goes as follows. If determinism is true, then our actions are the consequences of events that occurred in the remote past and the laws of nature. But it is not up to us what the laws of nature are or what went on in the past before we were born. Therefore, our present actions that are the consequences of these things are not up to us.

There are different compatibilist strategies to counter the Consequence Argument. Some argue that in a certain weak but sufficient sense it up to us what the laws of nature are or what went on in the remote past. Other compatibilists concede that the Consequence Argument shows that determinism implies that we do not the ability to act otherwise but then they argue that this ability is not necessary for us to have free will. They claim that being able to act freely is just a matter of being able to deliberate rationally on how to act in light of one's desires and not being under any physical constraints like being paralyzed or threatened at gunpoint. And these conditions can obviously be satisfied in a deterministic world.

One of the reasons that the question of free will has raised so much interest is that many people think that having free will is necessary for being morally responsible for one's actions. When we say that a person has done something wrong and blame her for what she did, we seem to imply that she should have refrained from performing this action. But what if she could not have acted otherwise? Harry Frankfurt famously argued that being able to act otherwise is not required for moral responsibility. Most of the philosophers that adhere to this idea are incompatibilists when it comes to free will and determinism but they believe that moral responsibility is compatible with determinism. John Martin Fischer, who takes this position, has named it *semi-compatibilism*.

But maybe our world is not deterministic. Some philosophers believe that free will is possible in an indeterministic universe. Many of these philosophers are incompatibilists who also believe that indeterminism is in fact true and that we do have free will. Such a position is commonly referred to as *libertarianism*. However, other philosophers think that indeterminism is just as much a threat to free will as determinism is, or even more so. They argue that if an agent's action was not determined by the past and the laws of nature it is indeed true that alternative courses of action were open for the agent but then it seems to be just a matter of chance which action the agent actually performed. So, it was not up to her and therefore she did not act freely. Incompatibilists who take this stance are sometimes called free will skeptics or free will pessimists since they believe that it is impossible for us to have free will.

Research in other fields than philosophy might in some cases provide findings that turn out be relevant to these philosophical questions about free will. For example, the question whether free will is compatible with determinism is a philosophical question, but the question whether determinism is true is one for the physicists to answer. Another example of potentially relevant research are studies of the neurophysiological workings of the human brain during the process of decision making. However, there is an ongoing dispute among philosophers about whether or not the results that have been generated in this area so far can be of any use in the endeavor to solve the philosophical problems of free will.

² See, for instance, Hofer (2010, Section 4.4).

SOR

The *sorites paradox*, or the *paradox of the heap*, is known since antiquity. The discovery is attributed to Eubulides of Miletus, 4th century BCE. The paradox is usually presented as an argument with two premises, e.g. as follows:

- (1)
 - (i) One million grains of sand, suitably arranged, form a heap.
 - (ii) If $n + 1$ grains of sand, suitably arranged, form a heap, then n grains of sand, suitably arranged, form a heap.
 - (iii) Hence, 1 grain of sand, suitable arranged, forms a heap.

This is a paradox, because both premises are intuitively true, the argument is intuitively valid, and yet the conclusion is intuitively false. Since, by common sense, it is clear that premise (i) is true and the conclusion false, what appears most suspect is premise (ii). But it also seems very plausible that if you take a grain away from a heap, what remains is still a heap: one grain of sand could not make the difference between being and not being a heap.

The apparent truth of premise (ii) has to do with the fact that the noun ‘heap’ is *vague*. In the philosophical discussion, the term ‘vague’ is technical, and does not denote simply any kind of lack of precision or specification. But pointing out that ‘heap’ is vague does not by itself solve the paradox.

Usually but not invariably, an account of vagueness includes three parts: (a) a definition or characterization of vagueness; (b) a theory that blocks the paradox; (c) an explanation of why we are disposed to accept the paradoxical reasoning. In the current state of the art, there are five main types of account:

- (1) According to *epistemicism*, premise (ii) is simply false: there *is* a sharp boundary between being and not being a heap, where one grain of sand does make a difference. The problem is that we cannot *know* where the boundary is. On one account, due to Tim Williamson, this is explained as follows: knowledge requires a margin of error, but knowledge of where exactly the boundary is would be knowledge without a margin of error, and is therefore impossible.
- (2) According to *supervaluationism*, developed by Bas Van Fraassen, Kit Fine, and Rosanna Keefe, the application of ‘heap’ to n grains of sand is clearly true for some large values of n , clearly false for some small values, and *indeterminate* for values in between. Now, you can make ‘heap’ into a *sharp*, i.e., non-vague, noun simply by *stipulating* that it be true, or that it be false, for the intermediate values. This can be done in different ways. A sentence is *supertrue* if it stays true regardless of the way in which ‘heap’ is made sharp, and *superfalse* if it stays false. Premise (ii) is superfalse, on this account, because it is false on every way of making ‘heap’ sharp.
- (3) According to *degree theories*, developed by Joseph A. Goguen, Kenton F. Machina and others, sentences are not just true or false, but have numerical values between 0 and 1, where 0 is complete falsity, and 1 complete truth. Any real number in between is an intermediate truth value. On this account, every *instance* of premise (ii) (such as ‘If 976.453 grains, suitably arranged form a heap, then 976.452 grains, suitably arranged, form a heap’) is *almost* completely true. But as we *combine* all these instances in reasoning, falsity *accumulates*, and so the degree of truth drops. Since the conclusion of the argument depends on the truth of *every* such instance, its own degree of truth is very low, in fact near 0.

- (4) According to *contextualism*, developed by Diana Raffman, Delia Graff Fara and others, about vagueness, the boundary of a vague expression, such as the noun ‘heap’, is highly sensitive to the *context of use*, and quickly *moves* with shifts of context. On some accounts, the context is psychological, on other accounts, it depends on the *conversation* rather than on the minds of the speakers. The main idea is that in every context, there *is* a sharp boundary of ‘heap’, but it is never *in focus* of the mind or the conversation. We can never *see* a sharp boundary, because the boundary is never where we look. Therefore, we tend to believe there isn’t any.
- (5) According to *gap* theories, a more recent trend, premise (ii) is in fact true. Vague expressions can only be used in semantically acceptable ways in contexts where the sorites-inducing element of their meaning does not *in fact* generate the paradox. If for some reason, for some intermediate heap size, there aren’t any heaps around of that size in the context, or if we can leave them outside what is semantically relevant in that context, then premise (ii) can be accepted as true, for the conclusion cannot then be derived. This is meant to explain why vague expressions can still be coherently used in almost every ordinary context.

Data and methods

In this study, the publication data comes from Web of Science. Web of Science involves seven citation indices, and we searched all of them in order to obtain publications on FW and SOR. For FW, the query $TS = (\text{“free will”})$ was used. This query corresponds to a topic search (TS) for the phrase “free will”. For SOR, we used the query $TS = (\text{sorites OR vagueness})$. From each of the two retrieved sets of publications, we extracted all publications of the document types *article*, *book*, *book chapter*, *proceedings paper* and *review*. For SOR, we further extracted each publication belonging to the Web of Science subject category *Philosophy*. This was done in order to exclude publications not relevant to the sorites/vagueness topic. We ended up with a set of 1,302 publications for FW, say P_{FW} , and a set of 377 publications for SOR, say P_{SOR} . Thus, the sample for FW is about 3.5 times higher than the sample for SOR. This fact should be kept in mind when the results are interpreted. For both FW and SOR, a great majority of the publications are articles: 1,167 of the FW publications (90 %) and 347 of the SOR publications (92 %) are of the document type *article*.

Development of the number of publications over time

In order to monitor the development of the number of publications over time, we assigned publications in P_{FW} and P_{SOR} to five-year time periods, based on the publication year of the publications. For FW, the periods are 1960–1964, ..., 2005–2009, whereas the periods for SOR are 1965–1969, ..., 2005–2009 (only a small number of the publications in P_{SOR} are published before 1965). The FW and SOR publications that do not belong to any of the time periods for the subdomains are not considered in this part of the study. The number of considered publications for SOR is 305. Regarding FW, cf. the next paragraph.

Since the problem of free will is discussed within several different faculties, like Humanities and Social Sciences, we decided to assign publications in P_{FW} to the four faculties Humanities, Medicine, Natural Sciences & Engineering and Social Sciences

(including Law).³ Thomson Reuters assigns journals, books and book chapters indexed in Web of Science to one or more subject categories, like *Computer Science*, *Hardware & Architecture* and *Ethics*. In the Web of Science records over, for instance, journal publications, one or more of these categories are represented (in the field *Web of Science Categories*). 1,021 of the Web of Science records that correspond to the 1,302 publications in P_{FW} are such that (a) the field in question is present in the record, and (b) the corresponding publication belongs to one of the time periods for FW. For each such record, the category expressions were extracted. Then we manually assigned each category to a faculty. For instance, *Philosophy* was mapped to Humanities. In a few records, the category *Multidisciplinary Sciences* is the only category represented. In these cases, we assigned the category to a faculty on the basis of information in the titles and abstracts of the records. These operations yielded, for a given publication, a list of one or more faculty expressions (a given expression can occur more than once in such a list). Then we applied fractionalization: for each publication P , each faculty expression E occurring in the list of P was assigned the fraction m/n , where m is the number of occurrences of E in the list and n the number of faculty expression occurrences in the list. For instance, for a publication with the list (“Humanities”, “Humanities”, “Social Sciences”), “Humanities” is assigned $2/3$, and “Social Sciences” is assigned $1/3$. Finally, for a given faculty and a given five-year period, the fractions for the faculty across the publications assigned to the time period were summed. The resulting sum is the number of fractionalized publications for the faculty with respect to the time period. Note that the number of (whole) publications for FW and a given time period is equal to the sum of the four faculty sums for the period.

Citation analysis

For each publication in P_{FW} and in P_{SOR} , we extracted the cited references (corresponding to cited publications) of the publication. The references were then standardized, i.e., different references referring to the same publication were mapped to the same reference. After standardization, and for FW, 34,214 unique cited references were obtained. The corresponding number for SOR was 4,751. Since we wanted to study, not only cited publications, but also cited authors and cited journals, we also extracted, for each publication in P_{FW} and P_{SOR} , the cited author names and the cited journal names of the publication. Both author names and journal names were standardized, in the sense indicated above. After standardization, 17,335 unique author names and 4,144 unique journal names were obtained for FW. The numbers for SOR were 1,948 and 447, respectively. In Web of Science, and with respect to cited references, only the first author of the cited publications is recorded. For fields where co-authorships are common, like physics, this limitation renders the cited author analysis problematic. For a field like philosophy, though, where a large share of the publications have exactly one author, the limitation is not a serious one.

Let the *citation frequency* for a cited reference (author name, journal name) x , with respect to P_{FW} (P_{SOR}), be the number of publications p in P_{FW} (P_{SOR}) such that p cites x .⁴ For FW (SOR), all cited references with a citation frequency ≥ 6 were selected, which yielded 372 (140) cited publications. For FW, and regarding author names and journal names, the 62 (50) most frequently cited author names (journal names) were selected. These two numbers correspond to a citation frequency ≥ 32 (34). For SOR, the 55 (50)

³ SOR is basically discussed within one faculty, namely Humanities.

⁴ We underline that the *citing* publications in this study (the publications in P_{FW} and P_{SOR}) exclusively are publications that occur as source items in Web of Science.

most frequently cited author names (journal names) were selected. These two numbers correspond to a citation frequency ≥ 17 (5).⁵

Let the *cocitation frequency* for two cited references (author names, journal names) x and y , with respect to P_{FW} (P_{SOR}), be the number of publications p in P_{FW} (P_{SOR}) such that p cites both x and y . Now, for FW (SOR), the cocitation frequencies for each pair of cited publications among the 372 (140) were calculated. Also for the selected author names and journal names, for both FW and SOR, cocitation frequencies were calculated.

We applied cluster analysis in the study. More precisely, we used the algorithm *Persson's Party Clustering* (PPC),⁶ which is similar to the single linkage clustering method (Everitt et al. 2001). This algorithm takes as input a list of pairs of objects, where the list is ordered descending by object-object similarity values. The algorithm reads this list, and takes one pair at a time. Clusters are allowed to form when at least three objects can be joined. Objects enter the clusters by their highest similarity values. Pairs of objects that are not clustered directly will be kept on a waiting list, which is searched every time a new pair is taken from the first list. The output of the algorithm is a list of clusters containing the objects. For mapping purposes, the clusters can be joined by finding pairs of objects, where the objects are from different clusters. In this study, we accept only one joining pair of objects for any two clusters: the pair with the highest similarity value (if the highest similarity value is equal to 0, the two clusters are not joined). For reason of graphical perspicuity, in the networks that follow, we have only indicated the links that have been active in the clustering process, in addition to the links between clusters.

For both FW and SOR, PPC was used to group cited publications, authors and journals, and cocitation frequencies were used as input similarity values. Singletons (clusters with exactly one object) were not generated.

Term analysis

A *term* in this study is a sequence of nouns and adjectives, ending with a noun. Terms in the Web of Science records corresponding to the publications in P_{FW} and P_{SOR} were extracted, where the record fields *Title*, *Abstract* and *Author Keywords* were used as term sources.⁷ However, some terms, mainly proper names, were standardized, in the sense indicated above, before extraction. The term extraction was performed by VOSviewer, a program for creating, visualizing and exploring bibliometric maps of science (van Eck and Waltman 2010). In the extraction process, VOSviewer uses natural language processing (NLP) techniques (van Eck and Waltman 2011).

Let the *occurrence frequency* for a term t , with respect to P_{FW} (P_{SOR}), be the number of publications p in P_{FW} (P_{SOR}) such that t occurs (with respect to the three fields mentioned above) in the record corresponding to p . For FW (SOR), 16,525 (2,775) unique terms were extracted. All terms with an occurrence frequency < 5 were excluded from the analysis. For each of the remaining terms, 865 for FW and 165 for SOR, VOSviewer calculated a relevance score, which indicates the specificity of the term in the text corpus given to VOSviewer as input. The top 50 % terms, with respect to relevance scores, were then selected by VOSviewer. Thus, 432 (82) terms were selected for FW (SOR).

⁵ For author names and journal names, we intended to select the 50 most cited names. Due to ties, however, the number of selected author names is greater than 50.

⁶ To our knowledge, PPC was first applied, with respect to research publications, by Persson (1994).

⁷ For both P_{FW} and P_{SOR} , all records of the publications have a title. However, there are records that lack either an abstract or author keywords, or both.

Let the *co-occurrence frequency* for two terms t and u , with respect to P_{FW} (P_{SOR}), be the number of publications p in P_{FW} (P_{SOR}) such that t and u both occur (with respect to the three fields mentioned above) in the record corresponding to p . For FW, we selected, after some tests, each pair of terms, among the most relevant terms selected by VOSviewer, with a co-occurrence frequency ≥ 4 for further analysis. The corresponding threshold for SOR was set to 3, again after some tests. After these selections, 178 (FW) and 52 (SOR) terms remained. PPC was then used to group the terms, with co-occurrence frequencies as input similarity values. Some singletons were generated, for both FW and SOR.

Concluding methodological remarks

In this study, similarity concerns thematic orientation in the context of (cited) publications, authors and journals, whereas similarity concerns semantic association in the context of terms. It should be clear from the two preceding subsections that cocitation frequency is used as measure of similarity in the former context, while co-occurrence frequency is used in the latter.

For the visualization of citations, cocitations, term occurrences, term co-occurrences and the outcome of the cluster analysis, we used Pajek, a program for analysis and visualization of networks (de Nooy et al. 2011). The Kamada-Kawai algorithm, implemented in Pajek, was used for automatic layout generation.

A large part of the data processing was done with the aid of Bibexcel, a toolbox for processing of bibliographic data (Persson et al. 2009). For instance, all PPC runs were performed by Bibexcel.

The philosophers Peter Pagin and Maria Svedberg wrote the introductions to SOR and FW (the preceding section), respectively. The two philosophers were given access to the results of the study after they wrote the introductions. Otherwise, they might have been influenced by the results. Pagin and Svedberg also interpreted the results. For the interpretation, the only data that were given to them were the data that occur in the figures and tables of the results section below.⁸ Pagin and Svedberg wrote parts of the results section, and they gave contributions to the final section, “Discussion and conclusions”. With regard to the results section, the text parts written by Pagin and Svedberg are in *italics*.

Cocitation frequencies and co-occurrence frequencies can be normalized with respect to individual citation and occurrence frequencies. We did some tests, where Salton’s cosine measure was applied (Salton and McGill 1983). However, according to the two philosophers, cluster solutions and networks based on normalized frequencies gave less valid representations of FW and SOR, compared to cluster solutions and networks based on raw frequencies. Therefore, we decided to use raw frequencies in this study.

Results

In this section, we report our findings. The first subsection treats the development of the number of publications on FW and SOR over time. In the second section, networks for FW based on citations, cocitations, term occurrences and term co-occurrences are visualized, whereas corresponding networks for SOR are visualized in the third section. We repeat that we have only indicated, for reason of graphical perspicuity, the links that have been active

⁸ Svedberg, however, requested information on which publications the small clusters of Fig. 3 contain. This information was given to her.

in the clustering process, in addition to the links between clusters. More links occur in the data than are visible in the networks.

The development of the number of publications on FW and SOR over time

In Fig. 1, the development of the number of FW publications over time is visualized. Four of the curves correspond to the four faculties, while the remaining curve concerns the total number of publications across all faculties. Note that regarding faculties, number of publications refers to fractionalized number of publications. From the period 1985–1989 to the period 2005–2009, and total number of publications, there is a pronounced increasing trend. We observe such a trend also for the faculty Humanities, regarding 1990–1994 to 2005–2009. With respect to the faculties Medicine, Natural Sciences & Engineering and Social Sciences, a clear growth from 2000–2004 to 2005–2009 can be observed.

In Fig. 2, the development of the number of SOR publications over time is visualized. As in the corresponding FW case, a pronounced increasing trend can be observed, here from the period 1990–1994 to the latest considered period 2005–2009.

The growth of publications over the years may, however, be inflated by the fact that a topic search in Web of Science records includes abstracts and keywords from 1991 and onwards. Moreover, it is known that Web of Science has broadened its coverage quite much during the last ten years. In order to analyze this problem, we limited the study to articles such that they (a) have the search terms in their titles, and (b) are published in journals that were active in 1980. In this way, two controlled sets of articles were obtained, one for FW, and one for SOR. There are 239 and 26 journals in the two corresponding

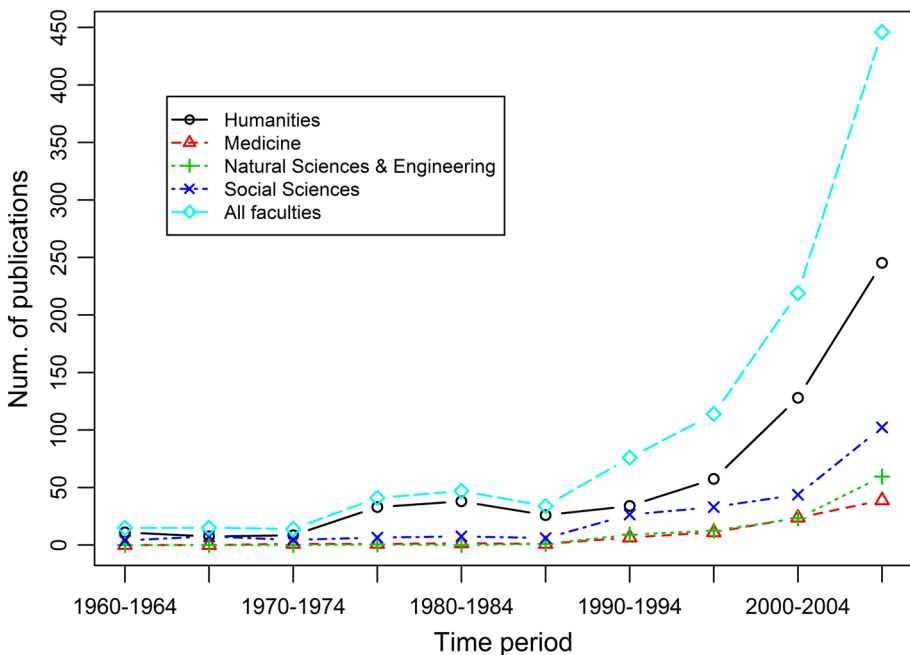


Fig. 1 Distribution of FW publications over faculties for 10 consecutive time periods, and the total number of publications across all faculties for each period. Number of considered publications: 1,021

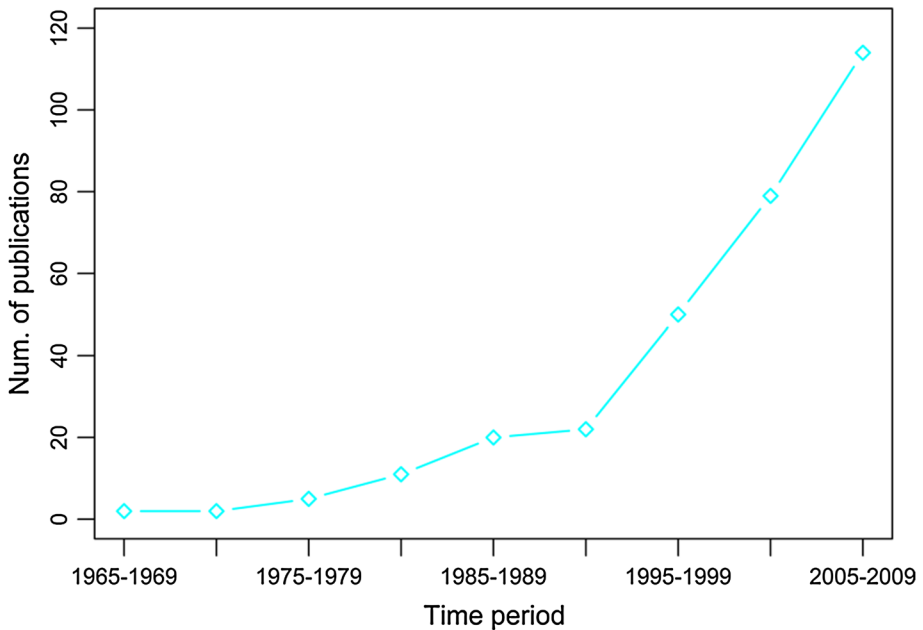


Fig. 2 Number of SOR publications over nine consecutive time periods. Number of considered publications: 305

control journal sets, respectively. For the set of FW articles, there was a 73 percent increase of the number of articles from the period 1990–1994 to the period 2005–2009. With regard to the faculty Humanities and the number of fractionalized articles, a 64 % increase was observed for the two periods. Increases were observed also for the other three faculties (Medicine, Natural Sciences & Engineering and Social Sciences). These increases cannot be considered as substantial, though, since the number of fractionalized articles that we here take into account is very small, for each of the three faculties. For the controlled set of SOR articles, the corresponding growth of articles was 106 percent. However, we should also take into account that the total number of articles published in the control journals might increase from the first to the second period. For FW (SOR), we observed a 23 (23) percent increase of the total number of articles in the 239 (26) control journals across the two periods. All in all, the findings suggest that the interest in the two subdomains has increased from the first half of the decade 1990–1999 to the second half of the decade 2000–2009, albeit not to the extent suggested by Figs. 1 and 2.

FW networks

Figure 3 shows a network of all cited (by the 1,302 FW publications) publications with a citation frequency ≥ 6 (372 cited publications). Node sizes correspond to the citation frequencies of the publications of the nodes, and a link between two nodes indicates that the two corresponding publications are cocited. Seven clusters, represented by different colors, were generated by PPC. For each cluster, the node for the most cited publication has a label that refers to the publication (the two largest clusters have more than one label). In Table 1, cluster numbers with corresponding colors and cluster sizes are given. The

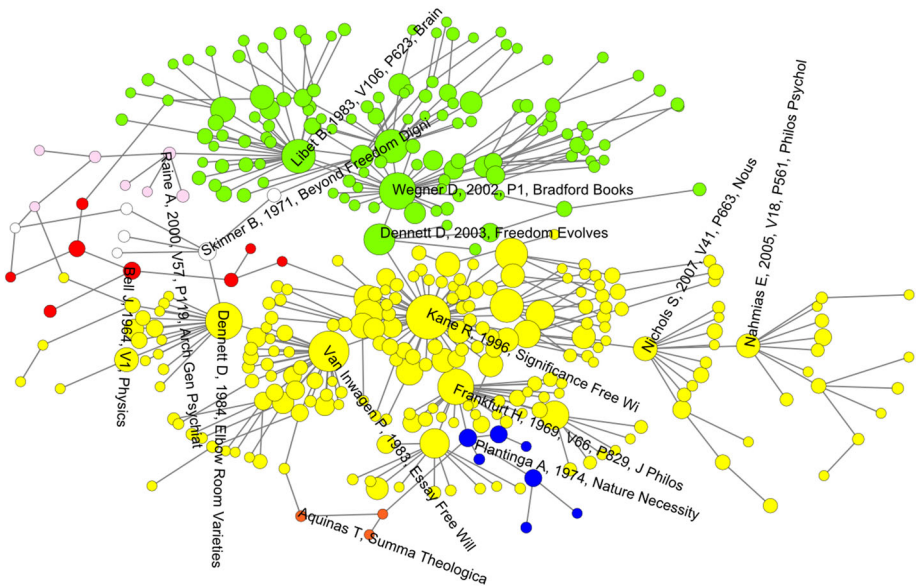


Fig. 3 Network of all cited publications with a citation frequency ≥ 6 (372 cited publications). Set of citing publications = P_{FW}

Table 1 Cluster number, cluster color and cluster size with respect to the PPC output for publications cited by FW publications

Cluster number	Cluster color	Cluster size (number of publications)
1	Yellow	215
2	Green	126
3	Red	8
4	Blue	7
5	Pink	7
6	White	6
7	Orange	3

network is clearly dominated by two large clusters, 1 and 2. These two clusters seem to be, in general, philosophy and neuroscience/brain science clusters, respectively.

The fact that publications in neuroscience/brain science form a cluster separate from the philosophy cluster means that cocitations within these two fields are more frequent than cocitations across the fields. There are, however, more links between cluster 1 and 2 than are visible in Fig. 3. Still, the separate clusters might indicate that even though the questions dealt with in neuroscience/brain science are relevant for philosophical issues on free will and vice versa, they are so only to a limited extent.

It is interesting to note that cluster 1 and 2 are connected by the philosopher Robert Kane in the philosophy cluster and the philosopher Daniel Dennett's second most cited publication (2003 Freedom Evolves) in the neuroscience cluster. In this book Dennett

deals with arguments on free will in both neuroscience and philosophy as well as in other areas. Dennett devotes one chapter to his criticism of Robert Kane’s libertarian idea on free will and that explains the fact that they have a high frequency of cocitations. But it is noteworthy that Dennett’s book belongs to the cluster dominated by neuroscience and not to the cluster dominated by philosophy publications. This may indicate that a majority of the philosophers have indeed considered Dennett’s discussion of arguments in neuroscience to be only marginally relevant to the philosophical problem of free will.

The largest cluster 1 gives a very representative image of the modern debate on free will in philosophy for the last 40–50 years. The four largest nodes in this cluster are publications by Robert Kane, Peter van Inwagen, Daniel Dennett and Harry Frankfurt and the ideas that they present have indeed had an enormous influence on the modern philosophical debate on free will. They discuss main issues such as compatibilism versus incompatibilism, the possibility of freedom under indeterminism and moral responsibility. In addition to these core parts of cluster 1, there is the right hand branch where we find publications like Nichols & Knobe (2007) and Nahmias et al. (2005). They are part of a discussion about the philosophical implications of ordinary people’s intuitions about free will and moral responsibility. In the debate between incompatibilists and compatibilists for example, philosophers on both sides have argued for their positions by referring to ordinary people’s intuitions in different particular cases. Some of the publications in this right-hand branch of cluster 1 report new experimental data on people’s intuitions and some deal with methodology problems and the philosophical relevance of surveys on folk intuitions.

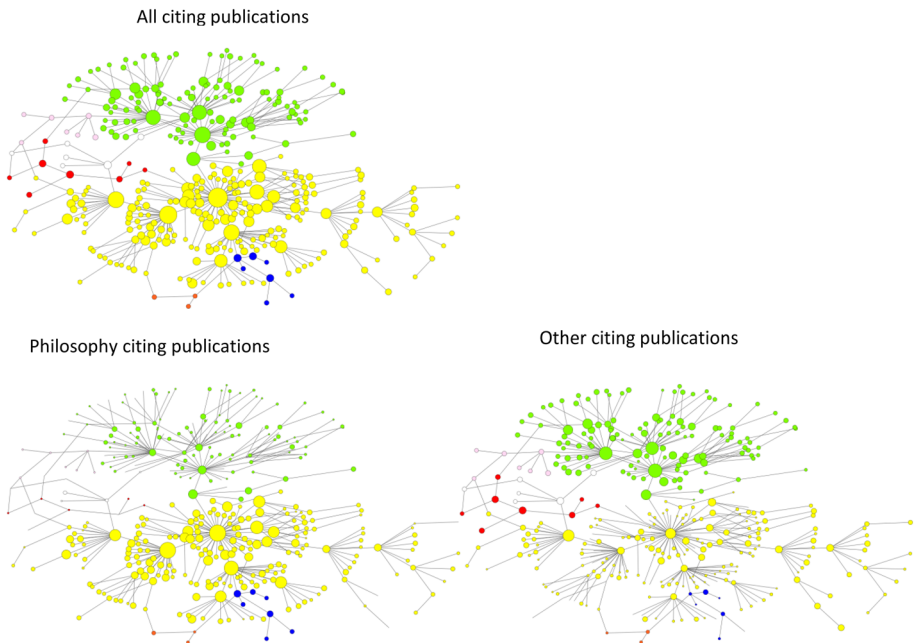


Fig. 4 Three networks of cited publications. The upper network is identical to the network given in Fig. 3. The lower left network is obtained from the upper network by only taken into account citations from philosophy publications in P_{FW} (512 publications). The lower right network is obtained from the upper network by only taken into account citations from non-philosophy publications in P_{FW} (790 publications)

We can also note that beside the main philosophy cluster 1, there are two much smaller clusters (4 and 7) that also consist of philosophical publications. Cluster 4 is connected to cluster 1 by Alvin Plantinga's book *The Nature of Necessity* which is on modalities and problems in the philosophy of religion. The rest of the publications in cluster 4 are modern theological classics. One reason why these publications are cited in the free will debate is that some of them deal with theological fatalism, which is the idea that God's infallible foreknowledge about everything that happens implies that every human act is necessary and therefore not free. The smallest philosophy cluster 7 contains one publication by Thomas Aquinas and two by Augustine.

Then we have cluster 3 with publications in physics and the small clusters 5 and 6 which both contain psychology related publications. The publications in cluster 5 are mainly on the effects of physical injuries and other dysfunctions in the brain on human behavior. For example, they discuss correlations between injuries in the prefrontal cortex and deficiencies in rational decision making and lack of impulse control over aggressive and violent behavior. Cluster 6 consists in large part of publications by B. F. Skinner on behaviorism. The fact that these psychology publications form separate clusters might indicate that even though the issues discussed in these publications are relevant to philosophical questions to some extent, these two fields are not that closely connected.

In general we can say that the formation of the different clusters in Fig. 3 is just what one should have expected. Free will is a research field that is quite diversified and problems related to the philosophical issues have been discussed and investigated in other areas than philosophy. And these areas are precisely those that have formed the separate clusters 2, 3, 5, and 6 in Fig. 3, i.e., neuroscience/brain science, physics and psychology.

We divided the set of FW publications, P_{FW} , into two disjoint subsets: the set of FW publications such that their corresponding records contain the term "Philosophy" in the field *Web of Science Category*, say P_{FW_P} , and the set of FW publications that do not satisfy this condition, say P_{FW_not-P} . The former set has 512 publications, and thus the latter set has 790 publications. The lower left network in Fig. 4 is obtained from the upper network, which is identical to the network given in Fig. 3, by only taking into account citations from publications in P_{FW_P} to the 372 publications. The lower right network is obtained from the upper network by only taking into account citations from publications in P_{FW_not-P} to the 372 publications. When only philosophy publications are considered as citing publications, nodes in cluster 2 (neuroscience/brain science) lose, in comparison with cluster 1 (philosophy), many of their citations. Conversely, when only non-philosophy publications are considered as citing publications, nodes in cluster 1 lose, in comparison with cluster 2, many of their citations. *This phenomenon was quite expected. What is noteworthy is that the nodes did not lose more citations than they did. This shows that there are indeed connections between philosophy and neuroscience/brain science when it comes to questions of free will. Some of the publications in cluster 2 concern studies of neural activities in the human brain during the process of decision making. There has been a discussion if these studies succeed in telling us whether or not our decisions to act in a certain way are determined by neural processes before we consciously intend to perform the action. There has also been a discussion about what implications the results from these studies might have for our ability to act freely.*

Further, the eight nodes in cluster 3 (physics) receive almost all their citations from non-philosophy publications. Note also that some nodes that are present in the upper network are absent in one of the lower networks. For example, the lowest node in cluster 3 is absent in the lower left network: the corresponding publication is no longer cited when only philosophy publications are considered as citing publications. This might seem a bit

Table 2 The 50 most frequently cited publications (citation frequency ≥ 20). The publications are ordered ascending by publication decade, and within the same decade descending by citation frequency. The rightmost column gives the cluster numbers for the publications

Cited publication	Decade	Citation frequency	Cluster number
Hobart R, 1934, V43, P1, Mind	–1959	23	1
Frankfurt H, 1969, V66, P829, J Philos	1960–1969	74	1
Strawson Peter F, 1962, V48, P1, P Brit Acad	1960–1969	32	1
Kornhube.Hh, 1965, V284, P1, Pflug Arch Ges Phys	1960–1969	22	2
Frankfurt H, 1971, V68, P5, J Philos	1970–1979	50	1
Van Inwagen P, 1975, V27, P185, Philos Stud	1970–1979	20	1
Van Inwagen P, 1983, Essay Free Will	1980–1989	93	1
Dennett D, 1984, Elbow Room Varieties	1980–1989	78	1
Libet B, 1983, V106, P623, Brain	1980–1989	65	2
Libet B, 1985, V8, P529, Behav Brain Sci	1980–1989	64	2
Strawson G, 1986, Freedom Belief	1980–1989	48	1
Frankfurt H, 1988, Importance What We C	1980–1989	34	1
Watson G, 1987, V96, P145, Mind	1980–1989	28	1
Watson G, 1982, Free Will	1980–1989	27	1
Kane R, 1985, Free Will Values	1980–1989	26	1
Nagel T, 1986, View Nowhere	1980–1989	22	1
Nozick R, 1981, Philos Explanations	1980–1989	20	1
Kane R, 1996, Significance Free Wi	1990–1999	113	1
Fischer John M, 1998, Responsibility Contr	1990–1999	60	1
Fischer J, 1994, Metaphysics Free Wil	1990–1999	47	1
Mele A, 1995, Autonomous Agents Se	1990–1999	44	1
Kane R, 1999, V96, P217, J Philos	1990–1999	43	1
Double R, 1991, Nonreality Free Will	1990–1999	36	1
Haggard P, 1999, V126, P128, Exp Brain Res	1990–1999	36	2
Ginet C, 1990, Action	1990–1999	34	1
Wolf S, 1990, Freedom Reason	1990–1999	33	1
Wallace R, 1994, Responsibility Moral	1990–1999	32	1
Dennett D, 1991, Consciousness Explai	1990–1999	28	2
Libet B, 1999, V6, P47, J Consciousness Stud	1990–1999	27	2
Wegner D, 1999, V54, P480, Am Psychol	1990–1999	26	2
Crick F, 1994, Astonishing Hypothes	1990–1999	22	2
Strawson G, 1994, V75, P5, Philos Stud	1990–1999	22	1
Widerker D, 1995, V104, P247, Philos Rev	1990–1999	21	1
Damasio A, 1994, Descartes Error	1990–1999	20	2
Wegner D, 2002, P1, Bradford Books	2000–2009	77	2
Pereboom D, 2001, Living Free Will	2000–2009	64	1
Dennett D, 2003, Freedom Evolves	2000–2009	55	2
Mele A, 2006, Free Will Luck	2000–2009	41	1
O'Connor T, 2000, Persons Causes Metap	2000–2009	41	1
Clarke R, 2003, Libertarian Accounts	2000–2009	33	1
Nichols S, 2007, V41, P663, Nous	2000–2009	33	1

Table 2 continued

Cited publication	Decade	Citation frequency	Cluster number
Soon C, 2008, V11, P543, Nat Neurosci	2000–2009	33	2
Nahmias E, 2005, V18, P561, Philos Psychol	2000–2009	32	1
Smilansky S, 2000, Free Will Illusion	2000–2009	31	1
Greene J, 2004, V359, P1775, Philos T R Soc B	2000–2009	30	2
Kane R, 2002, Oxford Hdb Free Will	2000–2009	29	1
Vohs K, 2008, V19, P49, Psychol Sci	2000–2009	25	2
Ekstrom L, 2000, Free Will	2000–2009	24	1
Libet B, 2004, Mind Time	2000–2009	21	2
Searle J, 2001, Rationality Action	2000–2009	20	2

surprising since these publications discuss quantum mechanics and hidden variable theories and these discussions are highly relevant for one of the core questions that philosophers are interested in: is our world deterministic or not? An explanation as to why these publications have a low frequency of citations from philosophical publications might be that the results that they present are indecisive.

The nodes in the psychology clusters 5 and 6, too, receive a major part of their citations from non-philosophy publications. The reason might be that in these publications we find studies that concern free will in a sense that is quite different from the ideas of free will that philosophers are interested in. More evidence for this hypothesis is to be found in Fig. 7, which shows a network of frequently occurring terms in the FW publications.

Table 2 gives the 50 most frequently cited publications (citation frequency ≥ 20). Note that we only consider publications in P_{FW} as citing publications. The cited publications are ordered ascending by publication decade, and within the same decade descending by citation frequency. *Here we can see that all of the 50 most cited publications belong to one of the two large clusters 1 and 2 in Fig. 3. 34 of the publications belong to the philosophy cluster 1 and among these there are no surprises. All the big names are there and no particularly influential publication is missing. Maybe it could be considered a bit remarkable that R. E. Hobart's publication from 1934 made it into the list. The three most cited publications are all written by philosophers (Robert Kane, Peter van Inwagen and Daniel Dennett). The fourth most cited publication is written by psychologist Daniel Wegner. Out of the 34 most cited philosophy publications there are only two that are written by women: Susan Wolf (1990) and Laura Ekstrom (2000). This correctly reflects the ratio of women to men in the philosophical fields of free will (and in philosophy in general) for the past 40–50 years.*

Another noteworthy fact is that of the 34 publications in the philosophy cluster there are 23 books (among these there are two anthologies) and 11 journal articles. In philosophy it is quite common to publish one's work in the form of an article but this study reveals that in spite of this trend, the majority of the most influential publications are books.

A network of the 62 most frequently cited (by the 1,302 FW publications) authors (citation frequency ≥ 32) is given in Fig. 5. Node sizes correspond to the citation frequencies of the authors of the nodes, and a link between two nodes indicates that the two corresponding authors are cocited. In this case, link widths correspond to cocitation frequencies. Exactly one cluster was generated by PPC.

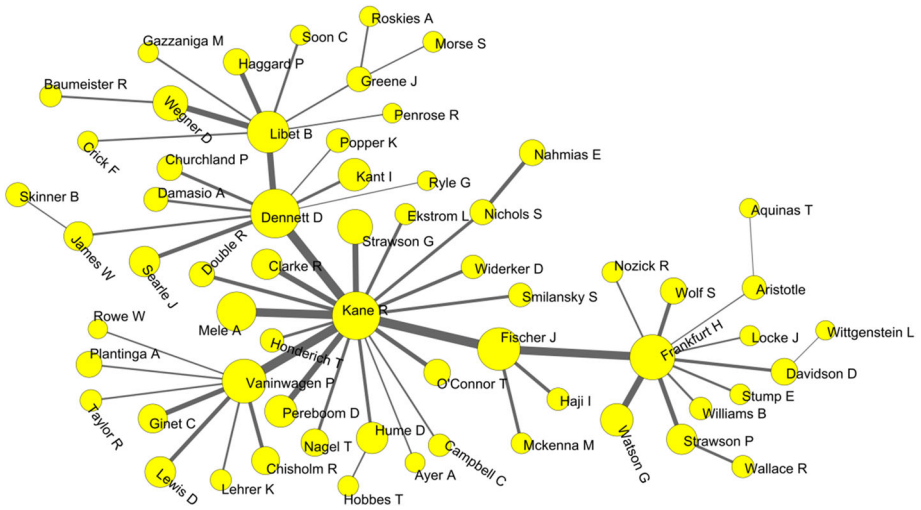


Fig. 5 Network of the 62 most frequently cited authors (citation frequency ≥ 32). Set of citing publications = P_{FW}

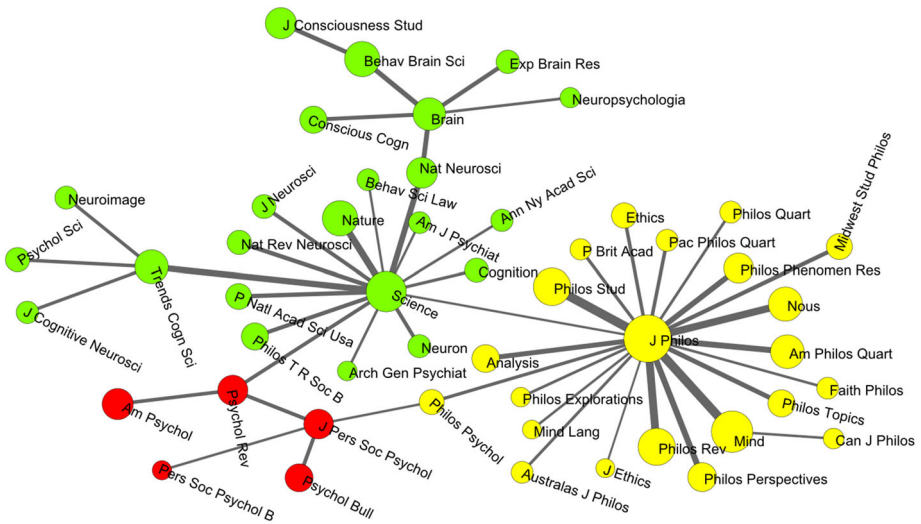


Fig. 6 Network of the 50 most frequently cited journals (citation frequency ≥ 34). Set of citing publications = P_{FW}

Here we can see that when we consider authors instead of publications historical names like Kant, Hobbes and Hume appear among the most frequently cited. This was not unexpected, since their views on free will have had a significant impact on some of the modern philosophical ideas. We can also note that a third influential woman within the philosophical field on free will appears among the most cited authors: Eleonore Stump.

The nodes that are linked to van Inwagen, Frankfurt, Fischer and Kane are those authors who are most prominent in the philosophical discussion on free will. The nodes

Table 3 Cluster number, cluster color and cluster size with respect to the PPC output for journals cited by FW publications

Cluster number	Cluster color	Cluster size (number of journals)
1	Yellow	22
2	Green	23
3	Red	5

that are linked to Libet are not part of the main philosophical debate and the nodes linked to Dennett are with a few exceptions not integral in the contemporary philosophical discussion of free will.

Figure 6 gives a network of the 50 most frequently cited (by the 1,302 FW publications) journals (citation frequency ≥ 34). Node sizes correspond to the citation frequencies of the journals of the nodes, and a link between two nodes indicates that the two corresponding journals are cocited. Link widths correspond to cocitation frequencies. PPC generated three clusters, represented by different colors. In Table 3, cluster numbers with corresponding colors and cluster sizes are given. *In cluster 1, we find well-known philosophy journals, like Journal of Philosophy (J Philos). Cluster 2 is dominated by neuroscience/brain science journals, whereas cluster 3 contains only psychology journals. This partition of journals shows again the diversity of the free will research field.*

A network of 164 terms occurring in (the records of) the 1,302 FW publications (occurrence frequency ≥ 5 ; co-occurrence frequency ≥ 4) is given in Fig. 7. Node sizes correspond to the occurrence frequencies of the terms of the nodes, and a link between two nodes indicates that the two corresponding terms co-occur. Link widths correspond to co-occurrence frequencies. PPC generated seven clusters (and some singletons, not represented in Fig. 7), represented by different colors. In Table 4, cluster numbers with corresponding colors and cluster sizes are given.

The terms in cluster 1 are those that commonly appear in philosophical writings on free will. This is however the case also for some of the terms in cluster 4, like “voluntary action” and “intentional action”. These terms are central in the philosophical debate on free will and the fact that they here appear linked to “neurobiology” and in the same cluster as “brain activity” and “participant” illustrates that studies of how our brains work during our decision making and acting are considered to be relevant for the philosophical questions about free actions.

In the main philosophy cluster 1 we can see that all the major positions in the modern philosophical debate on free will are represented in form of the names of their most prominent advocates. Van Inwagen is an avid incompatibilist with respect to both free will and moral responsibility. The compatibilist position is represented by Dennett. Frankfurt launched the idea that moral responsibility does not require free will in the sense of being able to act otherwise and Frankfurt’s idea was later embraced and developed by Fischer, who coined the term semi-compatibilism. Kane has presented one of the most influential libertarian theories on free will, i.e., free will under indeterminism. Finally skepticism about free will and moral responsible is here represented by Pereboom and Galen Strawson. However, the name Strawson might in some cases refer to Galen Strawson’s father P. F. Strawson who also is frequently referred to in publications on moral responsibility.

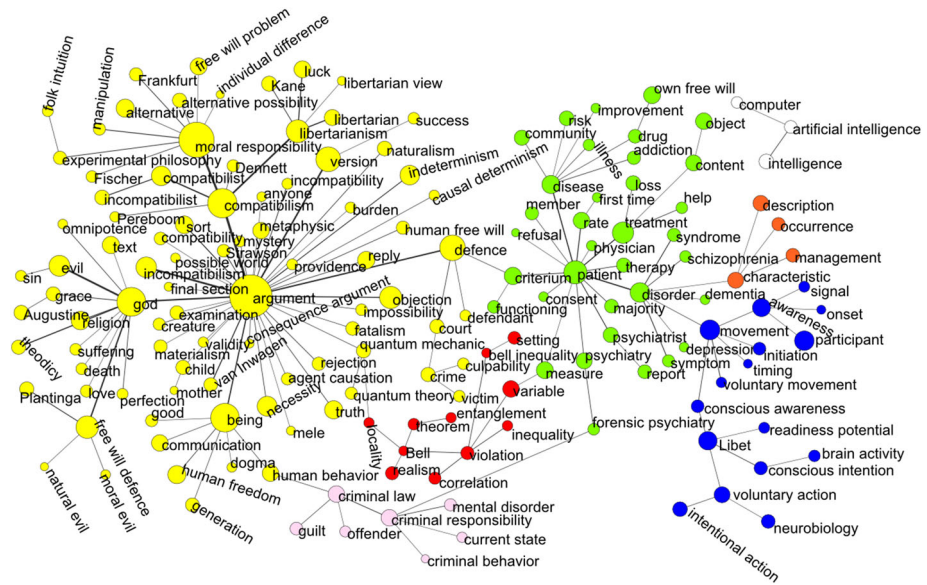


Fig. 7 Network of 164 terms occurring in the 1,302 FW publications (occurrence frequency ≥ 5 ; co-occurrence frequency ≥ 4)

Table 4 Cluster number, cluster color and cluster size with respect to the PPC output for terms occurring in the records of the FW publications

Cluster number	Cluster color	Cluster size (number of terms)
1	Yellow	88
2	Green	35
3	Red	11
4	Blue	16
5	Pink	7
6	White	3
7	Orange	4

The terms that appear in some of the smaller clusters suggest that some of the publications in the set of the 1,302 FW publications deal with a notion of free will that is quite different from the philosophical conceptions of free will. For example, in cluster 2 there are many terms that are related to psychiatric disorders, and in the discussion in that area the term free will might have a very loose or trivial meaning. This may also be the case in publications in the field of law and the terms in cluster 5, like “criminal responsibility” and “criminal law”, indicate that there are some law publications in our set.

SOR networks

Figure 8 shows a network of all cited (by the 377 SOR publications) publications with a citation frequency ≥ 6 (140 cited publications). Node sizes correspond to the citation frequencies of the publications of the nodes, and a link between two nodes indicates that the two corresponding publications are cocited. Exactly one cluster was generated by PPC.

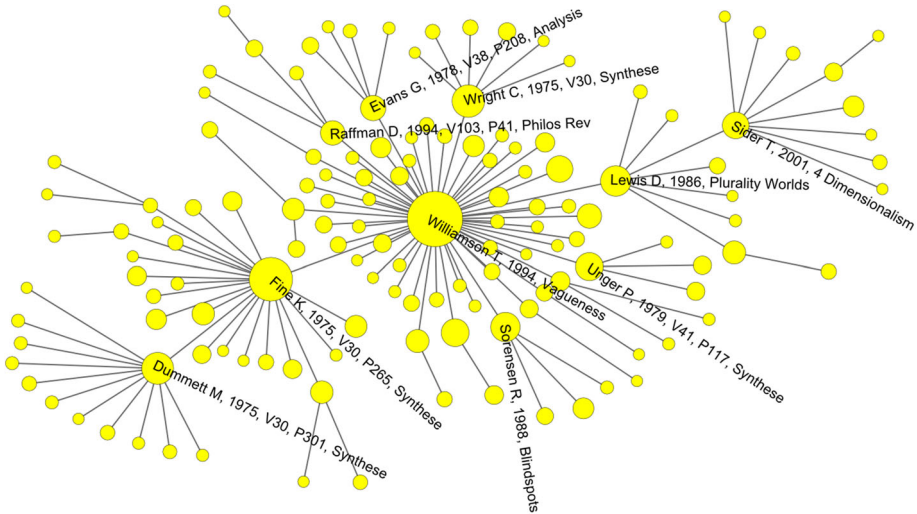


Fig. 8 Network of all cited publications with a citation frequency ≥ 6 (140 cited publications). Set of citing publications = P_{SOR}

Several highly cited publications are such that their corresponding nodes have labels referring to the publications.

The fact that a single cluster emerged from the data is not surprising, and would reflect the fact that the topic of vagueness and the sorites paradox is a fairly well delimited area. There is a technical sense of 'vagueness', closely related with the sorites paradox, that distinguishes vagueness from underspecification (or ignorance) more generally. Even though there is no generally agreed upon definition of vagueness, participants in the discussion agree on the delimitation of the set of vague expressions.

Table 5 gives the 46 most frequently cited publications (citation frequency ≥ 13). Note that we only consider publications in P_{SOR} as citing publications. The cited publications are ordered ascending by publication decade, and within the same decade descending by citation frequency. Since exactly one cluster was generated, the table does not have a cluster number column.

We see here a mix of publications in the vagueness area. The standard book on vagueness, Tim Williamson's book Vagueness, from 1994, clearly has the most citations, and the most influential technical account of vagueness, Kit Fine's supervaluationist account from 1975, comes in second. Otherwise, there is no particular theory or part of the area of vagueness that stands out in the list.

A network of the 55 most frequently cited (by the 377 SOR publications) authors (citation frequency ≥ 17) is given in Fig. 9. Node sizes correspond to the citation frequencies of the authors of the nodes, and a link between two nodes indicates that the two corresponding authors are cocited. Link widths correspond to cocitation frequencies. Exactly one cluster was generated by PPC.

Here we can discern three subgroups: a central part centered on works by Williamson, a left hand part centered on Fine's 1975 paper, and a right hand part centered on works by David Lewis.

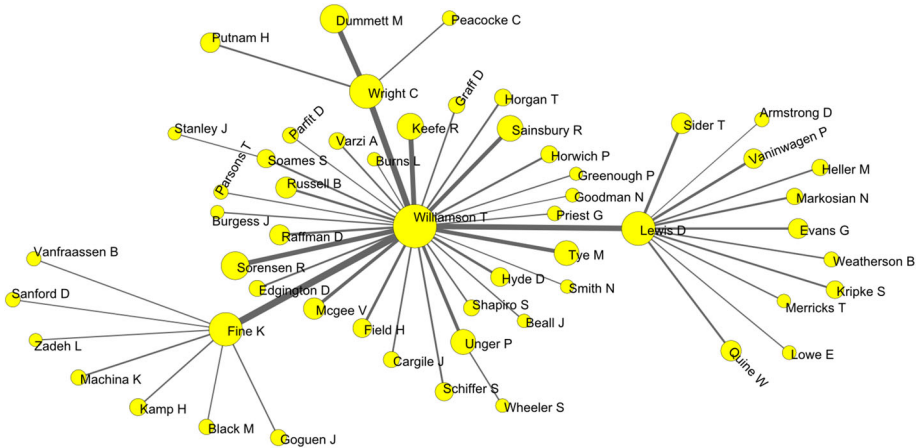
The left-hand part is concerned with advanced technical techniques for handling vagueness, and in particular supervaluationism and degree theories. The right-hand part is

Table 5 The 46 most frequently cited publications (citation frequency ≥ 13). The publications are ordered ascending by publication decade, and within the same decade descending by citation frequency

Cited publication	Decade	Citation frequency
Russell B, 1923, V1, P84, Aust J Philos	–1959	19
Goguen J, 1969, V19, P325, Synthese	1960–1969	23
Cargile J, 1969, V20, P193, Brit J Philos Sci	1960–1969	20
Quine W, 1960, Word Object	1960–1969	15
Fine K, 1975, V30, P265, Synthese	1970–1979	85
Wright C, 1975, V30, Synthese	1970–1979	48
Dummett M, 1975, V30, P301, Synthese	1970–1979	46
Unger P, 1979, V41, P117, Synthese	1970–1979	35
Evans G, 1978, V38, P208, Analysis	1970–1979	29
Wright C, 1976, Truth Meaning	1970–1979	22
Kamp H, 1975, Formal Semantics Nat	1970–1979	17
Machina K, 1976, V5, P47, J Philos Logic	1970–1979	17
Unger P, 1979, P177, Midw Studies Philos	1970–1979	15
Wheeler S, 1979, V41, Synthese	1970–1979	15
Campbell R, 1974, V26, P175, Philos Stud	1970–1979	13
Lewis D, 1986, Plurality Worlds	1980–1989	42
Sorensen R, 1988, Blindspots	1980–1989	39
Wright C, 1987, V15, P227, Philos Topics	1980–1989	24
Kamp H, 1981, Aspects Philos Logic	1980–1989	23
Unger P, 1980, V5, P411, Midwest Stud Philos	1980–1989	15
Putnam H, 1983, V19, P297, Erkenntnis	1980–1989	14
Lewis D, 1988, V48, P128, Analysis	1980–1989	13
Williamson T, 1994, Vagueness	1990–1999	139
Keefe R, 1997, Vagueness Reader	1990–1999	32
Raffman D, 1994, V103, P41, Philos Rev	1990–1999	27
Tye M, 1990, V99, P535, Mind	1990–1999	27
Van Inwagen P, 1990, Mat Beings	1990–1999	24
Soames S, 1999, Understanding Truth	1990–1999	22
Mcgee V, 1995, P203, So J Philos	1990–1999	21
Markosian N, 1998, V92, P211, Philos Stud	1990–1999	20
Hyde D, 1997, V106, P641, Mind	1990–1999	17
Lewis D, 1993, Ontology Causality M	1990–1999	17
Williamson T, 1992, V66, P145, Arist Soc P	1990–1999	16
Heller M, 1990, Ontology Phys Object	1990–1999	15
Tye M, 1994, V8, P189, Philos Perspectives	1990–1999	15
Horgan T, 1994, V8, P159, Philos Perspectives	1990–1999	13
Tappenden J, 1993, V90, P551, J Philos	1990–1999	13
Williamson T, 1999, V108, P127, Mind	1990–1999	13
Wright C, 1995, V33, P133, Southern J Philos	1990–1999	13
Keefe R, 2000, Theories Vagueness	2000–2009	35
Sider T, 2001, 4 Dimensionalism	2000–2009	32
Sorensen R, 2001, Vagueness Contradict	2000–2009	19

Table 5 continued

Cited publication	Decade	Citation frequency
Williamson T, 2000, Knowledge Its Limits	2000–2009	18
Shapiro S, 2006, Vagueness Context	2000–2009	16
Wright C, 2001, V110, P45, Mind	2000–2009	15
Stanley J, 2003, V63, P269, Analysis	2000–2009	13

**Fig. 9** Network of the 55 most frequently cited authors (citation frequency ≥ 17). Set of citing publications = P_{SOR}

clearly concerned mainly with ontic vagueness. In the center we find a mix works treating logical, semantic, epistemological and psychological aspects. Some of the positions are a bit surprising. One would for instance have expected Edgington to be lined up with the degree theorists in the left-hand subgroup.

Figure 10 shows a network of the 50 most frequently cited (by the 377 SOR publications) journals (citation frequency ≥ 5). Node sizes correspond to the citation frequencies of the journals of the nodes, and a link between two nodes indicates that the two corresponding journals are cocited. Link widths correspond to cocitation frequencies. Exactly one cluster was generated by PPC.

The journal cluster is unsurprising. *Mind* and *Synthese* are generally central and very prestigious journals. Especially the former would be central on just about any topic. Something that underlies their centrality especially in this area is that they publish both quite technical and completely untechnical papers, a mix that we do find in the sorites and vagueness areas.

A network of 44 terms occurring in the 377 SOR publications (occurrence frequency ≥ 5 ; co-occurrence frequency ≥ 3) is given in Fig. 11. Node sizes correspond to the occurrence frequencies of the terms of the nodes, and a link between two nodes indicates that the two corresponding terms co-occur. Link widths correspond to co-occurrence frequencies. Three clusters (and some singletons, not represented in Fig. 11), represented by

different colors, were generated by PPC. In Table 6, cluster numbers with corresponding colors and cluster sizes are given.

Here it is very clear that Cluster 1 is concerned the epistemology of vagueness and epistemicism as a theory of vagueness. It is equally clear that Cluster 2 is concerned with ontic vagueness, and the questions of how complex three-dimensional and four-dimensional objects are made up of parts. Cluster 3 is more of a mixed bag, but seems in part to concern properties of natural language, and its use, more generally.

Discussion and conclusions

The purpose of this study was to test the fruitfulness of advanced bibliometric methods for mapping subdomains in philosophy. The development of the number of publications on FW and SOR, the two subdomains treated in the study, over time was studied. We applied the cocitation approach to map the most cited publications, authors and journals, and we mapped frequently occurring terms, using a term co-occurrence approach.

Both FW and SOR show a strong increase of publications in Web of Science. When we decomposed the publications by faculty, we could see an increase of FW publications also in social sciences, medicine and natural sciences. The multidisciplinary character of FW research was reflected in the cocitation analysis and in the term co-occurrence analysis: we found clusters/groups of cocited publications, authors and journals, and of co-occurring terms, representing philosophy as well as non-philosophical fields, such as neuroscience and physics. The corresponding analyses of SOR publications displayed a structure consisting of research themes rather than fields.

Generally, for both FW and SOR, the most frequently cited publications in our material are books, rather than journal articles (Tables 2, 5), a finding that agrees with earlier bibliometric research on philosophy (e.g., Cullars 1998; Knieval and Kellsey 2005). Now, these highly cited books of the study, together with a considerable amount of other books, are not included in Web of Science. This is partly due to the fact that the two book indices of the database do not cover publications published earlier than 2005. Thus, the references from such books are not taken into account in the study. This is a minor concern for the part of the analysis that concerns cited publications. However, for the parts of the analysis that concern cited authors and cited journals, this limitation might be problematic: an excluded book might have several citations to a given author or journal. Therefore, it cannot be ruled out that the inclusion of non-Web of Science books would yield a somewhat different result.

Negative references, i.e., references going against the conclusions drawn in a sentence, are quite common in philosophy (e.g., Hellqvist 2010; Hyland 1999). Cocitation analysis builds on the idea that objects, like authors, are frequently cocited because they are similar. For two authors, for instance, a high cocitation frequency does not necessarily mean that they share the same standpoints, but it indicates that they are part of the same discourse. We do not believe, then, that negative references have a distorting effect on our results.

All in all, both philosophers involved in this work acknowledge the validity of the various networks presented. Most of the nodes, and the links between them, did reflect their conceptions of the subdomains FW and SOR. Even if the networks presented reflects the conceptions of the two philosophers, there are network properties that were surprising to them [cf. the philosopher comments on Figs. 4 (FW) and 9 (SOR)]. The philosophers found that the results of the two analyses give reason to some optimism about the prospects of bibliometrics as a tool in the sociology of scientific knowledge. Also Kreuzman (2001)

expresses optimism regarding bibliometric mapping and concludes that the cocitation approach is: “clearly valuable in going beyond the anecdotal evidence that there is a general lack of communication and interaction between philosophy of science and epistemology. Moreover, this author co-citation analysis provides a structural picture of the intellectual relations between these various philosophers which can provide the basis for further inquiry.”

To our knowledge, the domain expert approach used in this study is rarely employed in bibliometric studies. The point of the approach is to let domain experts assess whether the bibliometric mapping yields meaningful representations of the mapped fields. Given that the mapping is basically sound, studies like ours distinctly illuminate in figures and tables what subject experts already know (and to some extent, what such experts do not know) but what other scholars within the same superordinate discipline do not know. For instance, a philosopher that has not worked with the FW theme but wants to do so can read a paper like ours. In that way, the philosopher would obtain an immediate overview of FW, an overview that would be hard to obtain by scanning reference lists, searching databases like Google Scholar, and so on. Moreover, since the domain expert approach is used, the philosopher would obtain additional information by reading the comments of the experts.

In empirical natural language semantics, there is a tradition of looking at the distribution of words as correlated with linguistic meaning. One of the hypotheses in this tradition states that words with similar distributional properties have similar meanings (Rubenstein and Goodenough 1965; Schütze and Pederson 1995). This hypothesis has practical utility for data mining in texts, but is more problematic for studying meaning. A well-known problem is that it is hard to tell synonyms apart from antonyms: ‘rarely’ would have a distributional pattern similar to both ‘seldom’ and ‘often’. Nonetheless, there is a semantic dimension that is common to all three, the dimension of frequency.

As an analogy in bibliometrics, one could venture the hypothesis that publications, for instance, that treat the same topic or very similar topics would be frequently cocited. In practice, we would expect the match to be less than perfect. One source of error would be lack of knowledge. One publication might simply be much less known than another even if very similar in topic. Another error source is lack of understanding. It might simply take some time before a group of publications come to be recognized as being close in topic and perhaps as taking a common approach to a particular problem. A third possible source of error has to do with the sociology of citation: in some fields, including philosophy, there is a tendency to cite and discuss authors that already are famous, even though their arguments or theories might have been presented earlier by less famous people. It is hard to tell this apart from ignorance, though.

If we allow ourselves to assume that there are objective facts about topic similarity, and that at least in hindsight we could come to know them almost completely for a limited field, we could in principle use actual citation patterns to study the sociology of a particular field, by looking at the deviations in the cocitation patterns from real similarities. For instance, to what extent will cocitation between two publications emerge over time? In what cases is that best explained by a slow development of understanding that they are similar; in what cases by a recognition that one author has been given too little credit; and in what cases by the fact that the authors of the publications belonged to different scientific subcultures, which prevented cocitation during some initial years after publication?

One particular question concerning Fig. 9 in the present paper is that we do not see the contextualist authors as forming a distinct subcluster. Diana Raffman, Delia Graff (Fara), Stewart Shapiro, and Scott Soames are all contextualists about vagueness, and although differing between themselves are still fairly similar from an external point of view. Their

names all occur in the central subcluster, not connected to each other. This might be simply an artifact of the method used or the filters applied in the visual presentation, but it might also be because it has taken some time for a literature on contextualism as such to emerge. This in turn may have both epistemic and sociological explanations. Ideally, bibliometric studies could be a tool for the study of the emergence and spread of knowledge in a field, as well as of the obstacles to this spread.

In the study, we have analyzed two different and narrowly defined subdomains in philosophy. An interesting challenge for future research would be to start with a much broader set of philosophy publications to see if we could discover the same two subdomains as parts of a larger knowledge terrain.

Acknowledgments We would like to thank an anonymous referee for various useful comments.

References

- Baneyx, A. (2008). “Publish or Perish” as citation metrics used to analyze scientific output in the humanities: International case studies in economics, geography, social sciences, philosophy, and history. *Archivum Immunologiae Et Therapiae Experimentalis*, 56(6), 363–371.
- Cullars, J. M. (1998). Citation characteristics of English-language monographs in philosophy. *Library & Information Science Research*, 20(1), 41–68.
- de Nooy, W., Mrvar, A., & Batagelj, V. (2011). *Exploratory social network analysis with Pajek (Rev. and expanded 2nd ed.)*. Cambridge: Cambridge University Press.
- Dowell, E. (1999). Interdisciplinarity and new methodologies in art history: A citation analysis. *Art Documentation*, 18(1), 14–19.
- Everitt, B., Landau, S., & Leese, M. (2001). *Cluster Analysis* (4th ed.). London: Arnold.
- Hammarfelt, B. (2011a). citation analysis on the micro level: The example of Walter Benjamin’s Illuminations. *Journal of the American Society for Information Science and Technology*, 62(5), 819–830.
- Hammarfelt, B. (2011b). Interdisciplinarity and the intellectual base of literature studies: Citation analysis of highly cited monographs. *Scientometrics*, 86(3), 705–725.
- Hammarfelt, B. (2012). Harvesting footnotes in a rural field: Citation patterns in Swedish literary studies. *Journal of Documentation*, 68(4), 536–558.
- Hellqvist, B. (2010). Referencing in the humanities and its implications for citation analysis. *Journal of the American Society for Information Science and Technology*, 61(2), 310–318.
- Herubel, J. (1991). Philosophy dissertation bibliographies and citations in serials evaluation. *Serials Librarian*, 20(2–3), 65–73.
- Hoefer, C. (2010). Causal determinism. *The Stanford encyclopedia of philosophy*. 2012, from <http://plato.stanford.edu/entries/determinism-causal/>
- Hyland, K. (1999). Academic attribution: Citation and the construction of disciplinary knowledge. *Applied Linguistics*, 20(3), 341–367.
- Kabelka, G. (2012). The development of Lithuanian philosophy during 1960–2010: Volume, institutions, publications. *Problemos*, 81, 109–123.
- Knievel, J. E., & Kellsey, C. (2005). Citation analysis for collection development: A comparative study of eight humanities fields. *Library Quarterly*, 75(2), 142–168.
- Kreuzman, H. (2001). A co-citation analysis of representative authors in philosophy: Examining the relationship between epistemologists and philosophers of science. *Scientometrics*, 51(3), 525–539.
- Larivière, V., Archambault, E., Gingras, Y., & Vignola-Gagne, E. (2006). The place of serials in referencing practices: Comparing natural sciences and engineering with social sciences and humanities. *Journal of the American Society for Information Science and Technology*, 57(8), 997–1004.
- Leydesdorff, L., Hammarfelt, B., & Salah, A. (2011). The structure of the Arts & Humanities Citation Index: A mapping on the basis of aggregated citations among 1,157 journals. *Journal of the American Society for Information Science and Technology*, 62(12), 2414–2426.
- Leydesdorff, L., & Persson, O. (2010). Mapping the geography of science: Distribution patterns and networks of relations among cities and institutes. *Journal of the American Society for Information Science and Technology*, 61(8), 1622–1634.

- Leydesdorff, L., & Rafols, I. (2011). Local emergence and global diffusion of research technologies: An exploration of patterns of network formation. *Journal of the American Society for Information Science and Technology*, 62(5), 846–860.
- Leydesdorff, L., & Salah, A. A. A. (2010). Maps on the Basis of the Arts & Humanities Citation Index: The journals Leonardo and Art Journal versus “digital humanities” as a topic. *Journal of the American Society for Information Science and Technology*, 61(4), 787–801.
- Lindholm-Romantschuk, Y., & Warner, J. (1996). The role of monographs in scholarly communication: An empirical study of philosophy, sociology and economics. *Journal of Documentation*, 52(4), 389–404.
- Manana-Rodriguez, J., & Gimenez-Toledo, E. (2013). Scholarly publishing in social sciences and humanities, associated probabilities of belonging and its spectrum: a quantitative approach for the Spanish case. *Scientometrics*, 94(3), 893–910.
- Meyer, M., Grant, K., Morlacchi, P., & Weckowska, D. (2014). Triple Helix indicators as an emergent area of enquiry: A bibliometric perspective. *Scientometrics*, 99(1), 151–174.
- Persson, O. (1994). The intellectual base and research fronts of JASIS 1986–1990. *Journal of the American Society for Information Science*, 45(1), 31–38.
- Persson, O., Danell, R., & Schneider, J. (2009). How to use Bibexcel for various types of bibliometric analysis. In F. Åström, R. Danell, B. Larsen, & J. Schneider (Eds.), *Celebrating scholarly communication studies: A Festschrift for Olle Persson at his 60th birthday* (pp. 9–24). Leuven, Belgium: International Society for Scientometrics and Informetrics.
- Rubenstein, H., & Goodenough, J. (1965). Contextual correlates of synonymy. *Communications of the ACM*, 8(10), 627–633.
- Salton, G., & McGill, M. J. (1983). *Introduction to Modern Information Retrieval*. New York: McGraw-Hill.
- Schütze, H., & Pederson, J. O. (1995). *Information retrieval based on word senses*. Paper presented at the proceedings of the 4th annual symposium on document analysis and information retrieval.
- Shen, J. T., Yao, L. Y., Li, Y. P., Clarke, M., Wang, L., & Li, D. (2013). Visualizing the history of evidence-based medicine: A bibliometric analysis. *Journal of the American Society for Information Science and Technology*, 64(10), 2157–2172.
- Tatry, M. V., Fournier, D., Jeannequin, B., & Dosba, F. (2014). EU27 and USA leadership in fruit and vegetable research: A bibliometric study from 2000 to 2009. *Scientometrics*, 98(3), 2207–2222.
- Thompson, J. W. (2002). The death of the scholarly monograph in the humanities? Citation patterns in literary scholarship. *Libri*, 52(3), 121–136.
- van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538.
- van Eck, N. J., & Waltman, L. (2011). Text mining and visualization using VOSviewer. *ISSI Newsletter*, 7(3), 50–54.