The citation culture
Wouters, P.F.

Citation for published version (APA):
Wouters, P. F. (1999). The citation culture

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
Chapter 8

Representing science

8.1 Introduction

A recurring theme in the use of science and technology indicators, as well as in the construction of new ones, has been the interpretation of these indicators. Given the dependence on citation data of the majority of interesting science and technology indicators, a general citation theory should make the meaning of science and technology indicators more transparent. Hence the continuing call for a citation theory in scientometrics and, sometimes, in science policy. So far, such a theory has not yet been developed by the experts in the field. This study suggests an explanation for this. In the final chapter, the solution I propose to the problem of citation theory will be set out in more detail. I will furthermore try to use this study’s results to better understand the ways in which the sciences are represented, both in the process of knowledge production itself and in science policy.

8.2 Summary of the results so far

This study defines and analyzes the citation culture by theoretically distinguishing the citation from the reference. Chapter 1 argues that it makes sense to analyze the reference and the citation on their capacity as signs. Since the citation and the reference have different referents and are actually each other’s mirror image, it does not seem very wise to blur the analytical distinction between them. This distinction has moreover the advantage that the quest for a citation theory in scientometrics and the sociology of science splits into two different, analytically independent research problems: the patterns in the citing behaviour of scientists, social scientists and scholars in the humanities on the one hand, and the theoretical foundation of citation analysis on the other. As a consequence, the locus of the birth of the citation shifts from the scientist’s desk to the indexer’s office. The question “how are these citations produced?” becomes more relevant than it would have seemed if this study had stuck to the conventional wisdom that the distinction between reference and citation is merely technically interesting.

Chapter 2 and chapter 3 study the production process of the citation by de-
scribing, on the basis of documents and interviews, the creation and building of the first SCI. The distinction made in chapter 1 proves fruitful: Eugene Garfield and Joshua Lederberg turn out to have accomplished the translation of Shepard’s legally bound citation into a new sign of science. Chapter 2 and chapter 3 also clarify the otherwise puzzling rejection of the SCI by many scientists. This resistance was inherently present from the very beginning and did not need any misuse before emerging. This is difficult to explain if one assumes, as most authors on the topic have done, that citation analysis is a more or less “natural” or plausible extension or use of the citation norms in science and the social sciences.

Chapter 4 follows up on this historical analysis by sketching some consequences of SCI’s publication. It describes the way the SCI was received by sociologists of science on the one hand, and by the science of science movement on the other. It also relates, on the basis of original Russian and Ukrainian publications and interviews, how the SCI contributed to the early emergence of two different types of scientometrics in the Soviet Union, one oriented towards natural scientists and the other to science policy institutions. The chapter argues that the response to the SCI differed. The Mertonian sociologists of science — who together with Derek de Solla Price were the first to systematically study the potential of the SCI as a sociological research instrument in their graduate seminar at Columbia University — incorporated citation data in their studies where they seemed to fit in with other data and with their theory. The science of science movement saw the SCI as its dream coming true: science would finally be enabled to analyze itself. This was moreover to provide the foundation for a rational science policy. Thus, the SCI was used in two different settings, Derek de Solla Price being one of the few who played a role in both contexts. The “capture” of the science of science tradition by the upcoming citation culture (embodied in the SCI) resulted in the social science specialty of scientometrics. This also explains why scientometrics has had a distinct position in the whole of science studies from the very beginning, in a perhaps somewhat subtler way than the usual distinction between those people who hate numbers and those who love them.

The SCI did, however, not only give rise to the specialty of scientometrics, it also laid the foundation of a whole set of new indicators of science and technology. Chapter 5 tries to spell out precisely how the citation is used as a building block of the new indicator building. The chapter also tries to capitalize on the distinction made in chapter 1 by looking at how the two different dimensions of “citing” and “being cited” are interacting in the shaping of novel indicators. In this way, chapter 5 shows the analytical fruitfulness of the point of departure of this study in the deconstruction of indicators. By analyzing in detail the most strategic types of indicator, i.e. the ones that are dominant in citation analysis, the implicit claim of the chapter is that this deconstruction can be done in essentially the same way on any scientometric or bibliometric indicator. Not so surprisingly, the chapter also unveils the “heterogenous engineering” character of the whole enterprise. Whereas indicators tend to be judged on their consistency, the creation of the prevailing scientometric indicators turns out to be a far from consistent process. Although this certainly does not invalidate their use, it does elucidate their character as semiotic and socio-cognitive products of creative scientometricians.
Tinkering with indicators is also a characteristic of science policy, as chapter 6 has shown. In this domain, the demand for measuring scientific performance acted as a catalyst. Science funding bodies, sometimes confronted with the conflict between a lack of budgets and rising numbers of good research proposals, wished to know how effective their funding policies were. This was the context in which Lederberg, sitting in at an NIH meeting in 1957, recalled the article by Gene Garfield in *Science* some years before. Essentially the same problem created the first Dutch citation analyses in the physics funding body FOM. In its turn, this stimulated the Science Advisory Council to develop an explicit indicator policy in its drive to open up the bastion of science in the Netherlands. Ultimately, this led to one of the largest scientometric centres in the world at Leiden University.

This pattern, in chapter 6 retold with respect to Dutch science policy, is not unique to the Netherlands. For example, a closely related question, “How well do big laboratories perform?”, was the research problem which led to the highly controversial performance indicators developed at the Science Policy Research Unit Sussex University at the end of the 1970s. The problem how well state-funded chemistry laboratories were performing led to co-word analysis at the Paris Ecole des Mines in 1976. Apparently, different political contexts, different periods, and different actors gave rise to very similar patterns: the creation of specific performance indicators and related scientometric methodologies. Without a demand for science and technology indicators, scientometricians would not have been urged to mass produce indicators. Science and technology indicators would probably still have existed, but would not have acquired the independent status they enjoy nowadays. In short, whereas chapter 4 and chapter 5 deal with the push emanating from the new symbols embodied in the SCI, in chapter 6 the building up of a market for scientometric indicators has been sketched.

Chapter 7 has tried to take a quantitative look at the core journal of scientometrics, *Scientometrics*. The empirical question in chapter 7 is a simple one: how has the specialty developed if we take the first 25 years of its core journal as indicative? The analysis shows first of all that notwithstanding its being born out of the fusion of the scientistically inclined science of science and the SCI, scientometrics displays the characteristics of a social science, not of a natural science. In this sense, Price’s dream of a more objective type of sociology, resembling the natural sciences, does not seem to have come true. On the other hand, scientometrics’ position in the domain of information science positions the field in the area of the relatively “hard” social sciences, as perceived by their practitioners.

Chapter 7 also shows that scientometrics, as represented by its core journal, does seem to have specific properties as a specialty. Whereas the patterns of both the co-author relationships and the citation network hint at scientometrics as a specific domain of scientific activity, the analysis of the title words does not reveal specific patterns. The words simply clump together in one big cluster. Apparently, scientometricians all speak the same language. Yet, research groups do not collaborate very strongly. This can be explained by the features discussed in

---

1. Ben Martin, Interview, Brighton.
3. This conclusion contradicts Schubert & Maczelka (1993).
chapter 6 and chapter 5. Access to clean data is a strategic asset and is always 
jealously guarded, not least because of the abundance of contract research. Cor-
dial relationships do indeed exist between different research groups, but data and 
methods of using them are not freely available.

8.3 A hybrid specialty

Since the seventies, several models of specialty development have been proposed 
in science studies. Most of these concentrated on successful and homogeneous 
academic specialties. For this reason the received specialty models do not seem 
well suited to deal with the hybrid entity this study deals with. In the construc-
tivist turn within science studies, they have moreover been criticized for their 
naive way of demarcating specialties. It may be better to combine insights from 
the older specialty models with those of the later constructivist studies which 
zoom in on scientists’ heterogeneous engineering. This leads to a heterogeneous 
model in which the development can be represented as being driven by both a 
pull and a push mechanism.

The push force would be provided for by the unfolding of the citation culture 
after the birth of the SCI. Since 1964, the citation sign has been mass produced 
on a regular basis (the SCI is published four times a year). Because of its specific 
semiotic properties and its relationship with various citing cultures in science, 
this created the possibility of a host of interrelated indicators (chapter 5). This did 
not materialize of its own accord, though. Signs generally do not move by them-
selves. The citations needed symbolic actors (people) for whom these symbolic 
possibilities were also social and economic opportunities to perform research and 
thereby create or sustain careers. The pull force would be generated by science 
policy, the central theme of chapter 6. Since science policy has passed the stage at 
which it merely distributed money for research, it has encountered an annoying 
paradox. To judge the state of affairs in research, it depends on the very expertise 
it is supposed to steer and shape. The scientometric indicators seem to provide 
a solution: they enable expert judgement which is not dependent on the scien-
tists involved. This promise of solving science policy’s paradox created a novel 
market for quantitative science and technology indicators, a process studied in 
chapter 6.

Thus, the push and a pull mechanisms would interact with one another in 
the building up of science and technology indicators and in the shaping of the 
specialty of scientometrics.

8.4 Indicators as translators

8.4.1 Science as an information cycle

The implications of the emergence of both scientometrics and a set of intercon-
ected science and technology indicators can be drawn out by taking a closer
look at the process of knowledge production. The production of scientific knowledge can be pictured as a cyclical process in which certain inputs, like money and labour force, are translated into certain outputs, like scientific articles and knowledge claims. The various steps in this process have been analyzed in science studies, both at the micro level (Knorr-Cetina 1981, Latour & Woolgar 1986), at the meso level (Lemaine, MacLeod, Mulkay & Weingart 1976), and at the macro level (Jasanoff 1990, Latour 1984). An important feature is the constant evaluation of scientific knowledge. This assessment is used in the setting of new targets and the writing of new research proposals. Figure 8.1 on page 199 gives a schematic overview of this cycle. As can be seen, peer review in its various forms is central to this cycle. This is the reason for naming it “the peer review cycle”.

The peer review cycle itself is the product of a rather complex and convoluted history. It should not be seen as science in its purest form, since it has itself been heavily influenced by science policy considerations in the past (Cozzens, Healy, Rip & Ziman 1990). Nevertheless, it is widely felt to be one of the most distinctive features of science. If the description of this cycle is, rather arbitrarily, started with the writing of a research proposal, the second step is the proposal’s evaluation by those peers involved (using both scientific and extra-scientific criteria). The research is then carried out, possibly leading to submission of an article. A
second form of peer judgement, organized by the editor of a journal, produces, if positive, a publication. A third form of peer review is the assessment, at regular intervals, of research groups, university institutions and even the national contributions to specialties as a whole. These evaluations are again based on expert judgement. The results of these diverse evaluations contribute to the process of priority setting at various levels, leading to new proposals for research, institutional transformations or priority programs. The whole of peer review procedures is part of the reputational control system, as analyzed by Whitley (1984).

This is, of course, a rather general representation. The different stages in this cycle are not always executed and they may take various forms in different countries. It should moreover not be read as the claim that science proceeds in a single turn of the cycle. It is an abstract representation of a multitude of interlocking procedures. The scheme does represent, however, important properties of knowledge production. In representing science this way, “information” is taken as the entity that flows and as the substance that is translated in various forms during this cycle. It can easily be seen that this peer review cycle is in no way autonomous. For example, political priorities influence the step from evaluation to priority setting and to a growing extent also the peer judgement of research proposals by research councils. Monetary and economic arguments influence the overall science budget. And the peer judgement as well as the formulated scientific problems are contingent on culture at large. Nevertheless, the scheme shows the central position of scientific expertise. Whereas a large number of social and cultural factors influence the evaluation of science, the resulting quality judgement has to be justified in terms of the digestion of the cognitive products of science, as laid down in the scientific literature.

The dominant role of relevant domain-bound expertise is changed by the bibliometric indicators. Since their advent, a scientific publication can be measured by citation analysis or positioned by co-word analysis. Hence, the expert in the field is no longer the sole source of evaluation expertise. In terms of the information cycle representation, the bibliometric indicators appear in the form of an new, added cycle. This cycle processes information about the primary information cycle, i.e. the peer review cycle. Apparently, this secondary cycle produces and transforms information about information, meta-information. Hence, the secondary cycle is also an information cycle.

Contrary to the first cycle, the secondary cycle does have a distinct beginning. This is the consequence of its contingency on the primary information cycle. Its first step is the semiotic inversion of the reference into the citation (chapter 1). The citation indicators are the main pillars upon which maps of science are built. A slightly different translation process takes place in co-word analysis. In this case, the article is translated into a set of keywords (indexer determined keywords, selected title words or a selection from the abstract or full text). These keywords are subsequently used to construct co-word indicators. Again, maps of science can be created with the help of these indicators. It is also possible to combine co-word indicators and co-citation indicators to construct maps of science.

If citation analysis were simply a numerical mirror of peer review judgements, nothing much new would be happening. This is, however, not the case (chapter
5). The translation of references into citations creates additional degrees of freedom in handling citations. This is enhanced by the construction of more elaborate indicators and maps, all of which implicate numerous more or less arbitrary decisions. This does not mean that anything goes. It does mean, however, that functionally equivalent indicators can be made in several ways, as discussed in chapter 5.

Since the citation plays a crucial role in the second cycle, this cycle may be called “the citation cycle”. This needs some explanation. Derek de Solla Price was the first to formulate this concept (Price 1979). The cycle I am proposing is rather different from Price’s. The father of scientometrics wished to exhibit an interlocking metabolic complex of bibliometric (and scientometric) parameters in a comprehensive and integrated structure after the manner of the Nitrogen Cycle. (Price 1979, 621)

Price’s citation cycle follows the construction of the Science Citation Index itself and tries to quantify the relationships among such items as source authors, cited authors, source publications and citing and cited articles. Notwithstanding Price’s witty and clever text, it is not exactly clear what precisely cycles in his citation cycle. It looks more like a tourist streetcar route illustrating various aspects of the structure of science than like a metabolic cycle. Sentences like “For the next stage in the tour we enter the domain of citations” (Price 1979, 625), seem to indicate that this was exactly what Price intended. The cycle I am proposing is, on the contrary, a dynamic cycle in which information about the production of knowledge is processed and transformed. Its main results are representations of science: in the form of information on the performance of researchers, research institutions or other actors in terms of certain indicators; in the form of maps of

![Figure 8.2: The citation cycle](image-url)
8.4.2 Interactions between the cycles

The policy discussions as well as the validation of indicators by the experts in the fields to be assessed are based on translations of concepts in the domain of one cycle into concepts relating to the other. The following types of interaction between the cycles can be hypothesized to occur:

1. The indicators may influence the evaluation of science as such. This is for example the case whenever the citation frequency is used as a measure of scientific performance.

2. The indicators may in a more indirect way redefine the notion of quality in the realm of peer review judgements. This is, for example, the case whenever publication in a source journal of the SCI becomes an independent criterion in the assessment.

3. The maps of science (and other complicated science representations) may alter the evaluation process. These maps may for example influence the mental map of scientific experts judging a certain sub-specialty.

4. Scientific experts may be involved in the choice and validation of bibliometric indicators, and co-citation and co-word maps. As has been shown, this has become something of a routine procedure. Since the maps involved are highly sensitive to the thresholds used at different stages in the computational process, the measures used and their resulting maps may be fine-tuned to produce images that make sense to experts in the field. These validated maps may subsequently be used in follow-up analyses.

5. Scientists may also be involved in the validation of ranking lists. It is imaginable, for example, that experts can pinpoint anomalous ranking phenomena, if only because they know the people on the list (contrary to the scientometrician).

6. Scientific experts may also be directly involved in the construction of quantitative indicators. They may, for example, be used as a source of expert knowledge on specific features of the scientific literature involved.

All forms of interplay between the two information cycles have in common that they entail a translation of one type of science representation into another. In figure 8.3 on page 203, the upper right half represents the domain of the citation cycle with its formal representation of the scientific literature. The lower left half represents the domain of the peer review procedures with its stressing of the cognitive dimension of science. Since meta-information cannot in itself be

---

4Empirical research to what extent these interactions actually have occurred is outside of the scope of this study.
Figure 8.3: The cycle interactions

- Priority Setting
- Proposal
- Research
- Submitted Article
- Peer Review
- Evaluation
- Maps of Science
- Co-citation Indicators
- Co-word Indicators
- Keywords
- Citation Frequency
- Article
- Reference
- Citation
distinguished from information, the two cycles may interact easily. Science policy tends to promote this interplay. This means that the two cycles will tend to change the very foundations on which they are built. In other words, neither can easily be found by empirical means in their pure form. Consequently, influence of one on the other will tend to go unnoticed.

These interactions may influence the scientific system at all levels, from the individual scientist to the realm of science policy. First of all, the evaluation of science may now get two different kinds of input: one representing the conclusion of the field-specific expert and one representing scientometric expertise. The confrontation between these two forms of expertise has been a main feature of policy debates on measuring science. Because of this, the field-specific scientist no longer has a monopoly position in evaluating science. The two different forms of evaluation do not, by the way, have to be distinguishable in any clear-cut fashion. The various interactions mentioned above, will on the contrary promote the blending of the two perspectives. Nevertheless, they represent analytically different science representations. These differences create both the space and the need for negotiations and mutual validations between the expert opinion of the scientists and the scientometric expertise. Incidentally, the effort these translation processes cost is itself an indication that citation analysis is far from identical to peer review.

8.4.3 Credibility cycles

The citation cycle may also transform the way scientists earn recognition. These reward processes can also be represented with the help of cycles (Latour & Woolgar 1986, Knorr-Cetina 1981, Cozzens et al. 1990). Figure 8.4 on page 205 shows the credibility cycle as discussed in science studies (Latour & Woolgar 1986, Rip 1996). This sketch presumes an identity between the reference and the citation. Since this is no longer tenable, I propose an adapted cycle, shown in figure 8.5 on page 206, which takes the consequences of the citation cycle into account.

Since measuring performance indicators is based on fundamentally different expertise from judging the intellectual novelty of a paper, the credibility cycle bifurcates. A new loop is added, making the credibility cycle more complex. The appearance of scientometrics in these credibility cycles may indeed be the main cause of the need of an ethical consciousness in scientometrics: indicators have the potential to end as well as make careers. Of course, the extent to which scientific credibility is made dependent on quantitative indicators or on qualitative judgements may vary widely over different institutions and scientific cultures.

8.4.4 Implications of the citation cycle

The interactions mentioned above are all feasible, although empirical research of these interactions should clarify the extent to which they are actually changing the creation and use of scientific knowledge. In any case, scientometricians are often eager to include scientific experts in their validation work. This has been
The Credibility Cycle for a Scientist  

Figure 8.4: The classical credibility cycle

the case from the very beginning of scientometric mapping (Narin 1976). The same is true of the inclusion of scientists in the interpretation of maps of science and even in the construction of fine-grained indicators. The citation cycle also seems to influence the peer review cycle, although the record of evaluating specialties and disciplines gives a mixed picture. For example, the various evaluation committees in Dutch science policy have taken a rather different attitude towards the inclusion of quantitative and bibliometric indicators (Van der Meulen 1992). The inclusion of citation data in the assessment of the performance of individual scientists is even more controversial. In science policy, it often amounts to a “not done”, although citation data pop up in most evaluation exercises. The question of whether or not the emergence of the citation cycle has changed the notion of scientific quality in any fundamental way should be verified empirically.

Even if the concept of quality has not changed, however, the emergence of the citation cycle is a significant phenomenon in the scientific system. It constitutes an additional meta-information cycle. As a consequence, two analytically different science representations are produced. One is the domain of the expert in the

---

5Moed and Van Raan, Interview, Leiden 1995.
6Both the concept itself and the way it is used may vary and it may be difficult to distinguish between them.
field, the other the prerogative of the scientometrician. Since the regular evaluation of knowledge is a central axis around which the wheel of science spins, the citation cycle affects a vital part of scientific knowledge production. It would seem rather strange then, to dismiss science and technology indicators as being of only minor importance in the discussion on the research system’s future. Admittedly, information cycles are rather elusive entities. Given their nature as hypothetical constructions, their existence cannot be proved or disproved directly. Nevertheless, the citation cycle has led to such material phenomena as controversies over the use of indicators, an institute like the Leiden Centre for Science and Technology Studies, the specialty of scientometrics, the journal *Scientometrics* and lately even an international society, the International Society for Scientometrics and Informetrics ISSI. In other words, the citation cycle has affected both the institutional structure of science and the discourse on science evaluation.

### 8.5 Paradigmatic versus formalized representations

#### 8.5.1 Two representational domains

The citation culture has thus created two stable configurations: an intricate maze of science and technology indicators on the one hand, and the social science specialty of scientometrics on the other hand. Both have an important property in common: they represent science on the basis of formalized properties, be it in scientific literature, in its social or cognitive networks, or in its language. This is the main reason that the difference between the reference and the citation, made
in chapter 1, is relevant to better understanding the dynamics of science representations in science as well as in science policy. In the sociology of science as well as in scientometrics, the resistance towards citation or co-word analysis has generally been interpreted in terms of vested interests. The same is true of the analysis of the resistance of scientometricians against critics of their practice. The question then is, however, where do these interests come from? Why do actors define them this way? And what cultural resources are mobilized by them? I would like to suggest that this is where the symbolic representations of science hinted at in chapter 1 come in.

Most of the time, science representations may not be very important. After all, researchers just do their job. A vague notion of the state of their specialty may be sufficient, even if it were inconsistent and half-baked. As soon as science itself is the topic of research, however, it does matter. Vague notions of technical experts then have to be translated into some explicit assessment of the state of the art in specialty X. All actors dealing with science and technology then have to more consciously manipulate representations of these phenomena. Practising scientists maintain a specific representation of their specialty, and of science in general, based on their reading of the literature, their expert discussions at conferences, and their appreciation of the skills of their peers. Engineers also base the representation of the technology they create on their specialized knowledge and know-how. In other words, their representations are based on the substance of the matter, on the substantive content of the scientific knowledge or on the specific skills embodied in the technology.

The scientometric representations of science and technology, whether embodied in indicators or in a scientometric article, are very different: they purposely abstract from the specific substance and are only based on the formalized relationships between the entities (scientists, publications, texts or artefacts). This difference, which all actors seem to notice in passing, is, I propose, the consequence of the fundamental difference in the concept of information that the various representations of science and technology build on. In other words, the two types of representations, produced in the peer review cycle and the citation cycle, not only differ with respect to their use but also with respect to their constitution.

### 8.5.2 Two concepts of information

Two different theoretical concepts of information have been used in the information sciences. The first was proposed by Shannon & Weaver (1949). These authors provided a theoretical underpinning of the encoding of information and thereby wished to solve the technical problem of noise on the communication channel. In their view, information is a countable entity. The amount of information enclosed within a certain message is equal to the average number of digits required to code it. This concept of information has nothing to do with meaning, it is an abstract, dimensionless entity, comparable with the thermodynamic entity entropy (because the form of the mathematical equations is the same). It is related to the paradigm of formal logic in mathematics:

Logic, and by incorporation all of mathematics, was a game played with
meaningless tokens according to certain purely syntactic rules. All meaning had been purged. One had a mechanical, though permissive (we would now say non-deterministic), system about which various things could be proved. Thus progress was first made by walking away from all that seemed relevant to meaning and human symbols. We could call this the stage of formalized symbol-manipulation. This general attitude is well reflected in the development of information theory. It was pointed out time and again that Shannon had defined a system that was useful only for communication and selection, and which had nothing to do with meaning. Regrets were expressed that such a general name as ‘information theory’ had been given to the field, and attempts were made to rechristen it as ‘the theory of selective information’ — to no avail, of course. (Newell & Simon 1990, 112)

A diametrically opposed concept of information was proposed by Bateson (1980). His concept puts central what Shannon & Weaver’s (1949) filters away: meaning.

In 1979 anthropologist-philosopher Gregory Bateson offered another definition of “information”: “Any difference which makes a difference.” He said, “The map is not the territory, we’re told. Very well. What is it that gets from the territory to the map?” The cartographer draws in roads, rivers, elevations — things the map user is expected to care about. Data, signal (“news of a difference”) isn’t information until it means something or does something (“makes a difference”). The definition of information I kept hearing at the Media Lab was Bateson’s highly subjective one. That’s philosophically heartwarming, but it also turns out there’s a powerful tool kit lurking in the redefinition. (Brand 1987, 78–79)

As has been discussed in chapter 6, for a few decades after World War II science was a self-governing domain. For this reason, the experts in specialty X still play a dominant role in judging the state of affairs in specialty X. They regularly produce an explicit representation of their specialty, an overall judgement which is based on the specialty-bound paradigm (Hoyningen-Huene 1993, 131–162). This paradigm is formed from an assessment of the meaning of the various bodies of knowledge, practices, institutions and so forth. Hence, it is based on the differences that make a difference within the specialty. This “paradigmatic science representation”, to give it a name, is built of the stuff of Bateson’s information concept. Evidently, this is true even if the specialty is Shannon’s information science.

Since the seventies, science has increasingly been held accountable for its performance by science policy officials and the public at large. Sociologists of science, some historians of science and technology, as well as academically inclined science policy officials have been asked to study science in order to somehow steer science towards societal needs and desires. The study of specialty X by outsiders was thereby made possible and acceptable. Do these inquiries and representations from the outside occupy a different position? After all, historical and most sociological studies of science create different science representations from those that the specialists involved would produce. Yet, they also focus on
meaning, on the difference that makes a difference. For example, the development of twentieth-century physics in Germany may be a story about the evolution of the Weimar republic. School formation in British astronomy may be explained in terms of, for example, differing styles of the leaders involved (Edge & Mulkay 1976). It is no longer the meaning of specialty X to itself, but meaning nevertheless reigns. It is no longer the paradigm of specialty X which dominates, but a different one, the historian’s paradigm or the “disciplinary matrix” (Hoyningen-Huene 1993, 145) of the sociologist. Nevertheless, a paradigm does reign. These science representations from the outside are therefore also paradigmatic science representations based on meaning. They belong to the same class as the internal science representations discussed above.

This does not hold for the scientometric representation of science. Science and technology indicators create a “formalized representation” of science which initially neglects meaning. Of course, to interpret these representations one needs to attribute meaning again. The main point is, however, that this attribution of meaning can be postponed. This is crucial because it enables the manipulation of “meaningless” symbols, such as the citation. The sign citation is an entity like Shannon’s information concept and like entropy. Dimensionless, meaningless, countable. The formalized science representation therefore does not use (or maybe better, is not informed by) Bateson’s but by Shannon’s information concept. If this line of reasoning is valid, the most important aspect of scientometrics may not be its numerical character. Porter (1995) has pointed to the objectifying role of numbers in bureaucratic and policy contexts (see also Van der Meulen 1992). This is also true of science and technology indicators. Nevertheless, it should be noted that not all scientometric representations are numerical in nature. Science maps in particular are geometric representations, although they are based on computations. Moreover, paradigmatic representations may also contain computations and numerical assessments, whether made by the field-specific expert or a historian or sociologist of science. Therefore, although scientometrics is a metrics and therefore numerical in character, this is not its most important feature. The distinction between the formalized and the paradigmatic representations is the more fundamental distinction.

This conclusion also reaches back to chapter 1: in itself the semiosis of the sign citation (chapter 1) is not important. After all, every citation can easily be converted back to its reference, and one can also create symmetrical tokens denoting interactive processes. This kind of re-translation is even an important part of the production process of the SCI. The main reason that the semiosis of the citation and of the co-word are relevant, is the symbolic possibilities it creates to construct an encompassing formalized science representation that can compete at every level with paradigmatic science representations.

---

7 Ashmore (1989) made the distinction between studying science from the inside or from the outside, an important one in his analysis. In the perspective of this study, this distinction seems, however, less important than the distinction based on the types of information used.


9 This does not, in itself, explain why the formalized representations have become competitive. For this, one needs to look into the material interactions, i.e. the market for science indicators.
8.6 Indicator theories

The distinction between formalized and paradigmatic science representations can also be used to throw new light on the discussion on citation theories in the sociology of science and scientometrics. The search (in vain so far) for a definitive citation theory in scientometrics and the sociology of science has led to the regular production of reviews of theories of citation and citing (Narin 1976, Elkana, Lederberg, Merton, Thackray & Zuckerman 1979, Cozzens 1981, Cronin 1984, Edge 1977, Edge 1979, Gilbert & Woolgar 1974, Leydesdorff & Amsterdamska 1990, MacRoberts & MacRoberts 1989, Smith 1981). I will not repeat this here. The existing citation theories have tried to attribute meaning to the citation. Rephrased in more general terms, indicator theories try to attribute meaning to an individual indicator or set of indicators. In terms of the distinction between formalized and paradigmatic science representations, existing indicator theories in the sociology of science have grounded the meaning of an individual indicator (or set of indicators) in terms of a paradigmatic science representation. Using the criteria mentioned above, indicator theories can be grouped into four clusters, depending on the type of connection made between a set of indicators and a particular paradigmatic science representation: a “science of science” cluster, a “sociological” cluster, a “semiotic” cluster and an “information science” cluster of indicator theories. All theories put the relationship between a set of indicators and a specific paradigmatic representation in a central position.

- The science of science cluster

The science of science cluster is represented first and foremost in Derek de Solla Price’s work and in many publications in *Scientometrics*. The central notion is the reflexive use of science to measure itself (chapter 4):

I take the position that the workings of science in society show to a surprising degree the mechanistic and determinate qualities of science itself, and for this reason the quantitative social scientific investigation of science is rather more succesful and regular than other social scientific studies. It seems to me that one may have high hopes of an objective elucidation of the structure of the scientific research front, an automatic mapping of the fields in action, with their breakthroughs and their core researchers all evaluated and automatically signaled by citation analysis. (Price 1961, 194)

This approach has not produced a very detailed, nor a consistent theory about the citation. It was assumed that the regularities of science as a social activity would more or less directly produce a well-structured set of data that could subsequently be used to access the hidden laws of the growth of science. Finding a simplest description of these laws, comparable with the way physicists go about their job, was the primary motive in this work. This also meant that playing with the data and using the citations rather loosely was perfectly permissible and even encouraged.
• The sociological cluster

The science of science and the sociology of science both accepted citation frequency as a valid measure of scientific quality and as a sociometrically interesting link between authors or publications (chapter 4). The premise shared by everyone was that the reference and the citation were basically identical. This enabled an intuitive approach to the meaning of citation. Garfield, Price, Sher, and the sociologists of science at Columbia University all regularly referred to the act of citing to justify using citation data. For the latter group, this was especially appealing because the citation seemed to fit in very nicely with Merton’s norms of science (Merton 1973). Norman Kaplan provided for the first explicit Mertonian explanation of the citation in an NSF-funded project (Kaplan 1965). In this perspective the citation is seen as the embodiment of the giving of recognition to which the scientist is obliged. Since this leads to a symmetrical positioning of the citation, it means that, provided the normal statistical precautions have been taken, the number of citations received is directly proportional to the recognition acquired.

The Mertonian paradigm has produced a number of interpretations of the citation, each slightly different but all tied in with the central notion of science’s specific norms and rules. In more general terms, Mertonian indicator theories try to explain the citation by relating it to the citing behaviour of the scientist or scholar. This has also been done by competing sociological paradigms. For example, citations have been interpreted as a form of persuasion (Gilbert 1977). This study has also used a sociological indicator theory, albeit implicitly. In chapter 7 the citation patterns have been used as sociometric indicators, providing information about the relationships in scientometrics at group level. The attempt to translate citation patterns into behavioural characteristics is the most common approach in citation theories. The sociological cluster is therefore the largest cluster of citation theories.

The main reason that sociologists of science feel that this perspective has not produced the one encompassing citation theory, is the variety of behavioural characteristics underlying the citation patterns found in the literature. This is, however, the consequence of the semiotic inversion of the reference into the citation. This inversion is asymmetrical: whereas the references have very different characteristics (both textually and behaviourally), citations are all the same. The citation no longer betrays from what type of reference it was produced. This is why one should expect it to be difficult or even impossible to recreate the variety by citation analysis. Unless one re-translates the citation to the reference, that is, as is done in reference analysis. This is also why it is impossible to link the sign citation to some specific behavioural characteristic with respect to citing. Hence, although this type of research has delivered a reasonable amount of knowledge about citing cultures in science, in the quest for a citation theory it is a dead end.

• The semiotic cluster
In the semiotic cluster, created by the group of Michel Callon, not the citation but the co-word is the central “actant”. It is, like the citation, made from a word (usually either a keyword or a title word) by stripping every meaning from it. These co-words are subsequently used to construct maps that are claimed to be representative of fields of research and networks of relationships. This type of work was initially justified by criticizing the use of citation analysis, but in a thoroughly relativistic or a-sociological approach it has no need of this kind of justification. Different from the sociological cluster, these indicator theories do not try to relate to social behaviour but to semiotic networks. Intertextual relations between co-words are taken to reflect the development of networks of forces in science as well as the cognitive development of research. The signs live their own life and need no further justification.

- The information science cluster
  This cluster probably originated in the work of V. V. Nalimov in Russia (who saw science as a self-organizing process of information processing) in 1966, and ten years later in the work of Francis Narin (Narin 1976) in the United States, who took citation relations between journals as a sign of communication and the transmission of information. Parts of Garfield’s writing can also be seen as the attempt from an information science point of view to the citation, as can some of Derek de Solla Price’s articles. In general, the citation is related to processes of communication in science, or is seen as the upshot of an increasing complexity in these processes. Most of the time, the reference and the citation are taken to be identical. The distinctive feature of this group of theories is that it represents science as an information process, and abstracts itself from the substantive issues. In other words, it takes Shannon’s information concept as point of departure for all of science.

All four types of approach ground the formalized framework of indicators in some paradigmatic science representation. Having a formalized paradigm, the third and fourth cluster of indicator theories transform all of science into a formal domain and obliterate the difference between the paradigmatic and the formalized representations. In contrast, what the first two clusters have in common is that they ignore the relationships between the indicators in the realm of the formalized science representation itself. In them, the paradigmatic domain is predominant.

### 8.7 The rise of the formalized

Chapter 1 argued that because of the difference between the reference and the citation, the legitimation of citation analysis should be analytically distinguished from the study of citing behaviour in science. The results of the subsequent chapters enable a more general formulation of this conclusion. Because of the emergence of the formalized representations, stimulated by the creation of the SCI,
multiple relations have been created between the formalized and the paradigmatic representations of science (and technology). Every existing science or technology indicator theory is the embodiment of one possible type of relation within the domain of all possible relationships. Encompassing all this is not a sociological theory, but simply this proposal: to recognize the two different domains, to position each indicator theory accordingly, and to establish their interrelations.

In this sense, my proposal is also a theory, though a more abstract one: one could call this a proposal for a reflexive indicator theory. First, it is a theory about indicator theories because it explains how they can be related to one another and why the 30 year long quest for a citation theory has not been fruitful. Second, it is a theory about the indicators themselves, starting from the analytical distinction between the reference and the citation. Apparently, the two levels, usually kept strictly separate in science studies, go together seamlessly, an indication that the reflexivity issue can be fruitful indeed (Ashmore 1989).

As a theory about indicators, the reflexive citation theory borrows from all four clusters mentioned above. This proposal shares its reflexive character with the science of science cluster. The difference is in the appreciation of the nature of science. The sociological cluster has contributed to recognizing the variety in citing cultures upon which the reflexive citation theory builds further. The semiotic cluster and this proposal share the sensitivity to the way signs restructure and recreate reality. I have borrowed the use of Shannon’s concept of information from the information science cluster. The difference with both the semiotic and the information science cluster is that these two clusters, each in their own way, attempt to translate the whole of science into their formalized domain because their paradigm is ultimately a formalized one. In the reflexive citation theory the playing field is more level: both concepts of information play an important role and are often combined in the creation of representations of science.

Until the 1960s, formalized representations of science were subsidiary to paradigmatic ones. The regular publication of the SCI since 1964 has enabled the citation representation of science to lead a more independent life. Combining citation data with other formalized data (e.g. coword data, econometric data, computerized full-text analysis, logfile analysis of documents in cyberspace) is presently enabling more complex formalized science representations. It is now possible to combine various formalized and paradigmatic representations at the same time, for example to address policy relevant questions in more than one way. This is also the reason why different partial citation theories exist instead of one general one. It is both a matter of the analyst being able to adopt more than one perspective (Leydesdorff 1995), and the consequence of the relatively recent emergence of the domain of formalized representations in the scientific system.

Thus, the proposal for a reflexive indicator theory does not invalidate the use of specific indicator theories. It does, however, limit the claims to legitimation. The rift in science studies between scientometrics and more qualitative types of research is, I propose, not in the first place the reflection of the general cultural divide between number crunchers and innumerates. Rather it can be taken to represent, in the cognitive dimension, the difference between the formalized and

---

10 Which, in itself, does not mean that they should also exist.
the paradigmatic science representations. The fact that each citation analysis has to justify itself afresh in each study is the consequence of the existence of multiple relationships between the two domains. If the formalized domain had been tightly connected to some paradigmatic domain, this might have been different because every citation analysis would have had at least initially a clear-cut meaning.\footnote{Although it would not have prevented the possibility of several interpretations.}

The interaction between formalized and paradigmatic science representations also limit the extent to which the formalized domain might “take over”. First of all, people tend to attribute meaning to quantitative indicators or analyses. If these must circulate in society, they will therefore have to be translated into some paradigm. Second, the extent to which the formalized representations can alter the production of scientific knowledge or technological know-how is restricted by the embodiment of these practices. For example, it is not likely that publishing in science could completely be corrupted into a practice of collecting citations as a means in itself. Neither is it likely that the scientometric maps of science could replace the writing of reviews. Thus, if the formalized representations play any role, it will be more in the form of hybrid representations of science and technology than as “pure” formalized ones. Scientists, engineers and science policy officials will therefore have to handle both paradigmatic and formalized representations. Citing and publication patterns may be developed with an eye to both domains. The importance attributed to journals that are covered in the production of the SCI (“ISI journals”), is an indication of this trend.

The emerging new regime of electronic scientific publication may moreover impose a new relationship upon the paradigmatic and formalized science representations. “Citation” as well as “co-word” may acquire a novel function, at all levels. Since electronic publication seems able to produce and process a larger number of formal records, formalized representations can be expected to play an increasing role, both as an information retrieval tool and as forms of on-line quality control (crucial in an era of “accountable science”). Formalized representations may also increasingly appear in distributed form. For example, whereas the SCI had to be published separately from its sources and references, electronic publishing enables “live citation indexes” as part of the cited publications. Formalized representations may also be embodied in simulations of science. Being essentially algorithmic, the formalized domain may also become a point of departure for science (and scientometrics) by robots. It may for example become feasible to have computers construct new indicators in an evolutionary simulation of the scientific and technological system.

In short, the rise of a variety of formalized science representations, as well as an increasingly intimate interaction between formalized and paradigmatic science representations, may be a lasting result of the emergence of the citation culture in science.