Paul Wouters
Semiotics and Citations

In his essay "Semiotics and Evaluative Bibliometrics", Blaise Cronin discusses the implications of a tight link between research funding and citation analysis. He calls attention to a marketplace for a new species of sign—the citation (Cronin, 2000)—and suggests that semiotics offers the bibliometric community a suite of supra-disciplinary tools to "develop greater sensitivity to the variable symbolic significance of the signs they routinely manipulate and treat as quasi-objective indicators of quality, impact and esteem" (p. 450). This chapter aims to honor and further develop this perspective. I explore to what extent a semiotic approach may enable us to better understand the "constitutive effects" of performance indicators (Dahler-Larsen, 2013) and how this perspective can further contribute to a more inclusive attitude to the problem of citation theory (Nicolaisen, 2007).

Before recapitulating the key elements of Cronin's (2000) argument, we need to clarify what we mean by semiotics. It is usually presented as the science of signs (Eco, 1976), but can perhaps better be summarized as the systematic scholarly analysis of sign systems. Eco's (1976) definition of a sign is straightforward: "Everything that, on the grounds of a previously established social convention, can be taken as something standing for something else" (cf. Walker, 2014, p. 317). Semiotics has developed two different mainstream approaches, one developed by the father of structuralist linguistics, Ferdinand de Saussure (Day, 2005) (who used the term 'semiology'), and the other by the inventor of pragmatist philosophy, Charles Sanders Peirce (James, 1898). Saussure's thinking is firmly within the domain of the symbolic and focuses on structural motifs and differential inter-textual positions. Its analysis is confined to relations among signifiers and signifieds, and since signifieds are conceptual in nature, they too are part of the linguistic system. They are meanings, not actual objects of reference (Keane, 2003, p. 412). Peirce is interested in the link between signs and real world objects, based on his pragmatist philosophy. Consequently, both define the concept of "sign" rather differently. For de Saussure, the radical separation of the word and the world is fundamental (Irvine, 1989). Peirce, on the contrary, places as central the linkages between sign vehicles and real world objects. Given Cronin's interest in contributing to citation theory and making the link between citations and research behavior, it is not surprising that he adopted Peirce's approach.

For Cronin (2000), the usefulness of semiotics is evident: "What, after all, are references and citations if not signaling devices?" (Cronin, 2000, p. 440). However, this obvious fact is "all too easily overlooked", by which Cronin meant that he had overlooked it himself in his previous work on the citation process (Cronin,

1984). Many authors, inside of bibliometrics as well as outside of the field, still use the terms "reference" and "citation" interchangeably and see the difference between the two signs as trivial. With his essay, Cronin clarified this nontrivial distinction and drew the attention of the bibliometric community to the potential of semiotics to explain the differential roles of the reference, acknowledgement, in-text citation, and citation as measurement instruments, freed from the original citing context. In summary, the essay sought to demonstrate "how semiotics can contribute to the ongoing debate on the role and significance of citations in the primary communication system" (Cronin, 2000, p. 440).

Cronin was not the first to call attention to these important differences in bibliometric signs. Derek de Solla Price (1970) first made the distinction between the words *reference* and *citation*. Narin (1976, p. 3) an others (e.g., Nicolaisen & Frandsen, 2008; Egghe & Rousseau, 1990) followed suite. Egghe and Rousseau (1990) summarized the distinction as follows:

> If one wishes to be precise, one should distinguish between the notions 'reference' and 'citation'. If paper R contains a bibliographic note using and describing paper C, then R contains a reference to C and C has a citation from R (Price, 1970). Stated otherwise, a reference is the acknowledgement that one document gives to another, while a citation is the acknowledgement that one document receives from another. So, 'reference' is a backward-looking concept while 'citation' is a forward-looking one. Although most authors are not so precise in their usage of both terms, we agree with Price (1970) that using the words 'citation' and 'reference' interchangeably is a deplorable waste of a good technical term.
>
> (Egghe & Rousseau, 1990, p. 204)

It would be silly to claim that bibliometricians would not be aware of the differences in characteristics of the distribution of references versus citations. After all, the study of characteristics of these distributions belongs to the core of the field. Nevertheless, most scientometricians have tended to use the term 'citation' and 'reference' interchangeably. Not only is this in accordance with the meaning of the English word 'citation', it also facilitates the explanation of the number of citations as a measure of scholarly quality or impact. After all, it seems obvious that a work to which many researchers have referred is of more importance than a work that is hardly cited.

Wouters (1998) was the first to make the difference between the two signs the central point of departure for the development of citation theory. Building on this, Wouters (1999) concluded that a theory of referencing behavior should be seen as fundamentally distinct from a theory of evaluative bibliometrics. This was based on the statement that there is a fundamental distinction between *reference* and *citation*. By analyzing references and citations as different signs, they were essentially positioned as different objects. Their relation is one of descent: the *citation*

emerges in an act of "semiosis" (the creation of a novel sign) from the *reference*. This has an important implication: it is no longer the scientist who creates the citation. Its source lies in the citation index and the producer of that index is the creator of the sign *citation*.

Of course, this does not mean that the citation is created out of nothing, although it must be said that Wouters (1999) did not pay enough attention to this implication of his proposal. The raw materials of the citation signs are still the references and the link patterns among the references form constraints on the possible citation patterns. These constraints are relatively flexible. Herein is the expertise of the evaluative bibliometrician. Evaluative bibliometrics consists of the fine-tuned creation of different citation indicators from the pool of links between references and documents as well as from the links among references. Field-normalized citation indicators are an example of such fine-tuning. Some indicators are even combinations of citations and references, for example source-normalized indicators (Waltman & Eck, 2012) or other improved journal indicators (Nicolaisen & Frandsen, 2008). The need for technical expertise in evaluative bibliometrics to create and measure these indicators, whether this expertise is built into bibliometric algorithms or delivered by scientometricians, is itself an indication that the relationship between *reference* and *citation* is anything but self-evident or given.

Cronin (2000) developed this semiotic approach in much more detail than Wouters (1999), extending his earlier analysis of the citation process (Cronin, 1984). Moreover, Cronin (2000, p. 441) already expected forms of altmetrics:

> The web is giving rise to new modes of communication, representation, recommendation and invocation. The ways in which, and reasons why, individual researchers and scholars are mentioned, or linked to on the web, are multifaceted. It is conceivable that novel forms of signalling will evolve, which could also be used as indicators of cognitive or social influence within specific disciplines or communities of professional practice.

His essay emphasizes the polysemy of signs: "Multiple interpretations of references and their extra-textual import are possible" (Cronin, 2000, p. 440). He looked into the different relationships between the sign vehicles and their context and included acknowledgements into the analysis (which Wouters [1999] had ignored): "References and acknowledgements, along with citations, are first cousins in an extended family of scholarly signs" (Cronin, 2000, p. 441) (for more on the interrelationship of these signs, see Desrochers, Paul-Hus, and Larivière, this volume). The essay shows how technical semiotic analysis can clarify the difference between different scholarly signs. For example, the reference embedded in the text is a different sign from the full bibliographic reference at the end of the scholarly article.

1 Sign Triads

The "sign triad" developed by Peirce is central in this mode of analysis (Gluck, 1997): "This triad allows us to examine references and citations in terms of three common dimensions: (i) the carrier of meaning (sign-vehicle); (ii) the meaning or concept referred to (interpretant); and (iii) the object pointed to (referent)" (Cronin, 2000, p. 443).
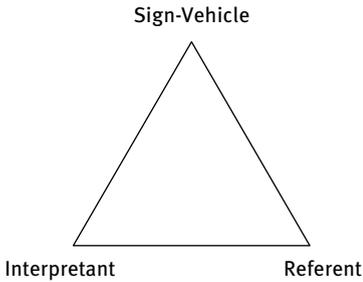
Sign-Vehicle

Interpretant   Referent

Fig. 1: Peirce's sign triad (after Gluck, [17]).

The triad makes visually clear the important differences between the embedded reference, acknowledgement statement, and individual citation as incorporated in the citation index and various forms of aggregated citations. The shape of the triad is identical, but the meaning of the three corners is different in different signs.
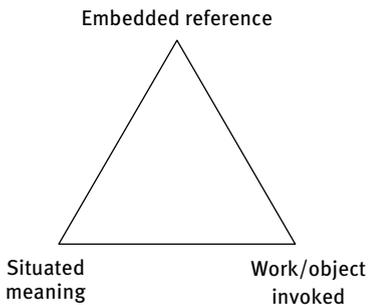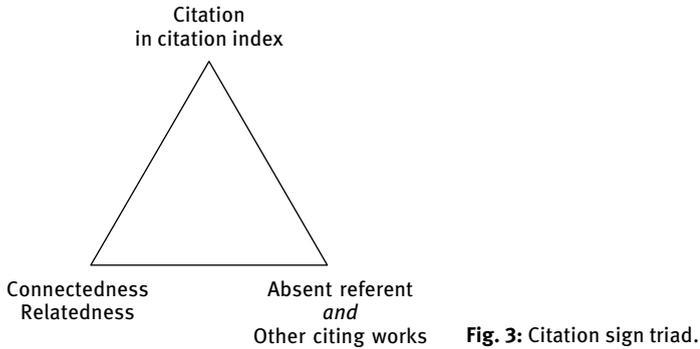
Embedded reference

Situated
meaning

Work/object
invoked

Fig. 2: Bibliographic reference sign triad.

In the case of the embedded reference, the sign vehicle is part of the citing text. Its interpretant is the concept flagged by the reference. This can be located in a specific part in the cited text (e.g., the methodology part), but it can also be more diffuse, such as when a complete book is cited. The embedded reference always

Citation
in citation index

Connectedness
Relatedness

Absent referent
and
Other citing works

Fig. 3: Citation sign triad.

has two different referents: the full reference in the bibliography (or footnote) of the citing text and the cited text.

In the case of the individual citation, the triad is quite different. First of all, the sign vehicle is no longer to be found in the citing text, but in the citation index. It is produced from the reference in the citing text, albeit not from the embedded reference but from the references as far as they are visible in the bibliography (footnotes are still poorly processed). It also has dual referents: it points back to its "parent text" and, in addition, all the texts that are listed by the citation index as having invoked it are its collective referent. The interpretant is also distinct from the interpretant of the embedded reference. Basically, it denotes inter-textual linking interpreted by Cronin (2000) as connectedness or relatedness.

Citation
aggregations

Impact
Esteem
Value

Authors

Fig. 4: Commodified citation sign triad.
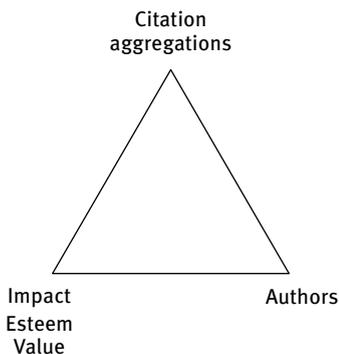
The aggregated citation sign is a more important sign than the individual citation since a lone citation does not make much difference. It is in the aggregation that the power of citation analysis manifests itself (see also Day, 2014). Again, the sign vehicle is located in the citation indexes, not in the form of a bibliographic string,

but rather as a calculated indicator. It can therefore also travel independently of the citation database proper, in the form of bibliometric reports or online bibliometric services. The referent relates to the unit of measurement at the specified level of aggregation. Cronin (2000) takes this to be the author, but of course this can also be the research group, the university, or the country as a whole. The interpretant is related to scientific impact as defined by the user of the citation analysis ("impact, esteem and/or value"). This can take subtly different formulations but the basic idea is that the more citations a work accrues, the bigger its scientific influence. In other words, the lower left corner of the commodified citation triad is the actual focus of evaluative bibliometrics.

Scientometricians have always examined the meaning and motivations underlying citations and the implications for citation analysis. However, since they have usually neglected to acknowledge the different configurations of the sign they were discussing, different realities have been collapsed into one another. In terms of the Peircean triads: the lower left corner of the embedded reference may be "situated meaning" and therefore able to explain why an author made the reference in the first place. Yet, in citation analysis, we are not dealing with embedded references but with the aggregated and commodified citations. Because the individual commodified citation is produced from the reference as listed in the bibliography, the original situated meaning of the embedded reference is lost in a two-step procedure. First, the embedded reference is transformed into the reference as listed in the bibliography. Then this list is inverted in the production of the citation index. Because the reference is decontextualized twice, the situated meaning cannot be aggregated and simply gets lost. This is why Wouters (1999) stressed that the resulting citation sign is in and of itself essentially meaningless. The application of citation analysis can no longer base itself on the original situated meaning of the embedded reference. A new theoretical foundation for citation analysis has to be created. This is the quest for the citation theory as conceived in the field of scientometrics (Leydesdorff, 1998; Nicolaisen, 2007).

According to Cronin (2000), the value of semiotics for the scientometric community lies primarily in its sensitivity to the variable interpretative possibilities of citations in the context of evaluative bibliometrics:

> Referencing is a complex phenomenon which can be analysed in terms of a set of sign systems (...). Furthermore, referencing and citation behaviours vary within and between disciplines, such that blanket criticism is misplaced.
>
> (Cronin, 2000, p. 445)

In addition, the processes of production and consumption of the relevant signs can be studied, which can be the basis for the study of the variation in the mean-

ing attached to these signs. In this context, Cronin pointed to studies showing that referencing behavior is governed by epistemic and social norms and values demonstrating that patterns in references are not completely arbitrary (in a debate with the McRoberts critique of evaluative bibliometrics as basically meaningless (MacRoberts & MacRoberts, 1996; MacRoberts & MacRoberts, 2009)). In other words, Cronin (2000) addressed "the meaning of the citation" as the key problem for citation theory.

Cronin (2000) acknowledged that it might become possible, especially in the context of the UK national research assessment procedures (Research Assessment Exercise (RAE), now called Research Excellence Framework (REF), see http://www.ref.ac.uk/) for the symbolic citation exchange processes to have material consequences:

> More specifically, if an individual's, department's or university's ability to amass symbolic capital of this kind were to become the critical determinant of future research funding and career advancement, then it would not be difficult to imagine distortions creeping into the system, as players devised recruitment, publication, collaboration and citation harvesting stratagems to accelerate and maximise the accrual of symbolic capital.
>
> (Cronin, 2000, p. 450)

Note the strong moral orientation of this fairly accurate prediction of the future—"distortions creeping into the system" imply a negative view towards these developments. Moreover, the semiotic analysis that Cronin (2000) advocates to understand the meaning of the citation is not mobilized to understand the material implications of the increased use of citations in the reputational system. They merely appear as an afterthought.

This conception of semiotics as relevant for scientometrics is therefore by definition restricted to the classical Mertonian approach in which the relationship between sign systems and behavior is central. References and citations are seen as representations of real world relationships. Because of this focus, the analysis proposed by Cronin (2000) and Wouters (1999) remains firmly in the symbolic domain. Their semiotic citation analysis is completely intra- and inter-textual and the material world of knowledge production does not have a direct reference in this semiotic storyline.

As a consequence, it is not entirely clear what can actually be done with this refined perspective on the signs reference and citation. The mechanisms through which linguistic signs might affect the material world of knowledge creation are not specified. It is especially difficult to understand how the subtle difference between an embedded reference and a citation that is part of a citation index leads to the conclusion that the number of citations one gets would not be a more or less direct consequence of the choice of scientists to refer to ones work. Most people

simply keep speaking about references and citations as more or less the same phenomenon. The details of the linguistic links seem no more than that: details that are only interesting for the person interested in the particularities of semiotic or linguistic theory. The semiotic approach was largely ignored by most researchers, including scientometricians dealing with citation analysis. It did impact the more theoretical discourse about citation theory, but even here its acceptance seems to have been limited to a small group of scholars. For example, in his comprehensive and well-documented review, Nicolaisen (2007) seems to conflate semantic and semiotic approaches.

At the same time, however, the practical application of citation analysis in the context of evaluation has taken off. Although the vast majority of "evaluation moments" are not recorded in a systematic way, it is clear that citation-based indicators have become available on a much larger scale than the inventors of the Science Citation Index could have predicted (Wouters, 2014). These indicators are used by managers as well as by researchers. It would be far too simplistic to assume that they are only management tools for which researchers can claim innocence. They have settled themselves firmly in the fabric of science and scholarship. In order to understand why and how this happened, we need an extended theoretical framework that is able to understand at a more fundamental level how the real world and semiotic signs like the citation and the reference actually interact.

2 Material Semiotics

This type of framework has already been developed, albeit outside of bibliometrics, and is called material semiotics. Material semiotics is not so much a unified theory as a family of attitudes and dispositions. They are relevant for the debate about the foundations of evaluative bibliometrics because they question the very distinction between the material and the symbolic. In order to understand how the sign system of the citation, produced by the citation indexer from the raw materials of the sign system of the reference delivered to the indexer by the research community, interacts with the material production of knowledge, we need an integrated way of thinking about, and acting with, both materiality and symbolism. The argument is succinctly formulated by the anthropologist Webb Keane:

> Efforts to bring theories of the sign into a full, robust articulation with accounts of human actions, self-consciousness, and social power are still commonly hampered by certain assumptions built into the lineage that runs from Saussure to post-structuralism. They tend still to demand that we divide our attention and choose between ideas and things. The re-

sult is that even those who would study "things" too often turn them either into expressions or communications of "ideas," or relegate those ideas to an epiphenomenal domain. Those who would study "ideas" too often treat the associated material forms as transparent, taking their consequentiality to be suspect, and, at times, imputing implausible powers to human desires to impose meaning on the world. And this divide seems to give rise to what is still a common, if ill-informed, perception among social analysts, that "semiotics" is a species of idealism.

(Keane, 2003, p. 410)

Keane is interested in the practical embodiments of semiotic ideologies in representational economies. This work is particularly relevant because research evaluation can be interpreted as an important moment in the exchange of reputation in the political economy of science and scholarship. He wants to draw attention to "the dynamic interconnections among different modes of signification at play within a particular historical and social formation" (Keane, 2003, p. 410). His research shows that how people handle and value material goods may be implicated in how they use and interpret words, and vice versa, reflecting certain underlying assumptions about the world and the beings that inhabit it (Keane, 2003, p. 410). His goal is to "open up social analysis to the historicity and social power of material things without reducing them either to being only vehicles of meaning, on the one hand, or ultimate determinants, on the other" (Keane, 2003, p. 411). He turns back to Peircean semiotics (in this sense his work is perhaps more linguistic in character than that of material semiotics in science and technology studies) because of its promise to overcome the radical sign-world dichotomy which is so characteristic for Saussurean analysis. Keane (2003) identifies two aspects of Peircean analysis as particularly important.

First, it is processual: "signs give rise to new signs, in an unending process of signification. This is important because, viewed sociologically, it can be taken to entail sociability, struggle, historicity, and contingency" (Keane, 2003, p. 413). It is striking how well this quote summarizes what has happened to citation indicators.

Second, it pays considerable attention to "the range of relationships not only between signifier (sign) and signified (interpretant) but between both of those and (possible) objects of signification" (Keane, 2003, p. 411). Keane argues that the nature of those relations between signifier and objects of signification needs to be grounded in the dynamics of the social relations. If we recall how Cronin (2000) used a triad to understand the difference between different forms of references and citations, it is clear that the anthropologist and the information scientist have found common ground, possibly without being aware of each other's existence. Keane's goal to "recognize how the cited materiality of signification is not just a factor for the sign interpreter but gives rise to and transforms modalities of ac-

tion and subjectivity" (Keane, 2003, p. 411) resonates with Cronin (2000)'s goal to "develop greater sensitivity to the variable symbolic significance of the signs they routinely manipulate and treat as quasi-objective indicators of quality, impact and esteem."

Building on various critiques of modernist and realist philosophy of science, material semiotics has developed particularly strongly in the field of science and technology studies to understand how science is able to create new worlds and how the interactions between "representations" and "real objects" can be analyzed (Latour & Woolgar, 1986; Latour, 1988; Haraway, 1991; Berg & Mol, 1998; Mol, 2002; Law, 2004; Luukkonen, 1997). Material semiotics is a rather fundamental and radical alternative to the dominant epistemology in European and American thought.

Material semiotics is radical in that it refuses to accept the separation between epistemology and ontology. It is not interested in the conditions for knowing (a central problem in classical epistemology) but in the ways objects are handled in practice and consequently in the performative nature of knowledge (Mol, 2002, p. 5). Science and scholarship do not analyze and represent a reality "out-there", but engage in the creation of new worlds "in-here" (Law, 2004, pp. 54–55). This does not mean, inter alia, that material semiotics denies the existence of reality. On the contrary, it pays far more attention to how realities emerge than mainstream social science tends to do. Its practitioners are also more sensitive to the role of technology and materiality in social processes than our usual sociologist or psychologist. In actor network theory, one of the main embodiments of material semiotics, the concept of *construction* is central (Latour & Woolgar, 1986). The key idea is that science creates new objects that are unstable and contested in their infancy, but then gets hardened into facts. Latour calls this the Janus face of science: Science in the making still recognizes the uncertain nature of scientific facts, but science in the classroom discusses facts as if they are part of a stable nature (Latour, 1987). Mol (2002) prefers to speak of the *enactment* of realities rather than of their *construction*:

> The term 'construction' was used to get across the view that objects have no fixed and given identities, but gradually come into being. During their unstable childhoods their identities tend to be highly contested, volatile, open to transformation. But once they have grown up objects are taken to be stabilized.
>
> (Mol, 2002, p. 42)

She does not adopt this notion of stabilization; in her perspective a certain fluidity is a stable feature of reality:

> ... the idea that objects might not just gradually acquire an identity that they then hold on to has been pushed aside, or complemented, by this new idea. That maintaining the identity of objects requires a continuing effort. That over time they may change. (...) If an object is real this is because it is part of a practice. It is a reality *enacted*.
>
> (Mol, 2002, pp. 43–44)

This has an important consequence which is relevant to citation theory: reality is multiple rather than singular. In her study of the treatment of atherosclerosis[1] in a Dutch hospital, Mol shows that it is not simply one disease reality that patients and doctors are dealing with. The disease turns out to be quite different objects. Atherosclerosis in the walking therapy session is distinct from atherosclerosis under the microscope, and different again from atherosclerosis as operated on by the surgeon. The reality of the disease does not precede the diagnosis and treatment, but is intertwined with them. In the diagnosis and treatment, the disease gets its specific form that defines it (Mol, 2002, p. 96–97). We are not speaking of different perspectives on one underlying reality here—the material interactions are different in the ontological sense. Nor are they disconnected realities, as they can be present within a single patient:

> It is one of the great miracles of hospital life: there are different atheroscleroses in the hospital but despite the differences between them they are connected. Atherosclerosis enacted is more than one - but less than many. *The body multiple* is not fragmented. Even if it is multiple, it also hangs together. The question to be asked, then, is how this is achieved.
>
> (Mol, 2002, p. 55)

The question of how actors (human and non-human) achieve this hanging together dominates Mol's (2002) work. She exposes the variety of strategies by which reality is made coherent in the practical interaction between humans and objects. Making the multiple character of reality invisible is an aspect of these strategies.

This perspective diverges fundamentally from the way we commonly think about both our social and our physical reality. The recognition of reality as multiple, rather than singular, has fundamental implications for the role of methodologies in the social sciences. These consequences have been explored by Law (2004). He suggests the concept of a *fractional object*: "We are in a world where bodies, or organisations, or machines are more than one and less than many. In a *world* that is more than one and less than many. Somewhere in between" (Law, 2004, p. 62). According to Law, there are three options in ontology. The first op-

---

1 Atherosclerosis is a disease that leads to thickened walls of the arteries, see http://en.wikipedia.org/wiki/Atherosclerosis

tion is to insist on singularity of the world. This means that those who perceive the world differently from ourselves are simply wrong. In social science, there is one best methodology to understand the world. The second option is to insist on pluralism and "the irreducibility of worlds, of knowledge, of ethical sensibilities, or of political preferences, to one another" (Law, 2004, p. 63). This is the relativist attitude. A large number of methodological choices in social science are now allowed, but the price we pay is that they are incommensurable. We have no criteria to decide whether a particular approach is better than another one. In contrast with these two options, Law advocates the third option which is in-between. Like Mol (2002), Law (2004) considers the world as fundamentally multiple. It is only through our active cohering of different practices and parts of realities that the world can develop as a coherent reality. This has an important consequence. The philosophical question on the nature of reality transforms into a political choice about how to live. Ontology is no longer in the first place a matter of *discovering* the real nature of reality, but a matter of making political choices as to which realities should be *created*. In terms of Law's and Mol's analysis: it is a matter of ontological politics. As a consequence, there is no general blueprint for social science, no generally valid methodology:

> There *is* no general world and there *are* no general rules. Instead there are only specific and enacted overlaps between provisionally congealed realities that have to be crafted in a way that responds to and produces particular versions of the good that can only ever travel so far. The general, then, disappears, along with the universal. The idea of the universal transportability of universal knowledge was always a chimera. But if the universal disappears then so too does the local - for the local is a subset of the general. Instead we are left with situated enactments and sets of particle connections, and it is to those that we owe our heterogeneous responsibilities.
>
> (Law, 2004, p. 155)

3 Semiotic Citation Theories

So what do these new developments in philosophy and material semiotics mean for the application of semiotics to the domain of citation theory? The step we need to take, which neither Cronin (2000) nor Wouters (1999) took, is to understand that the sign is not a representation of reality, however distorted, but an object in the real world of knowledge making. The differences between the embedded reference, the reference in the bibliography, the disembedded citation which is the in-between half-product hidden in the deep recesses of the citation index machinery, and the citation as part of the citation index, are not differences of *representation* of one and the same real world object, but *ontological* differences

between objects. Each of them is a partially different but related object. This first conceptual step is not easy because it requires us to distance ourselves from the usual mainstream American/European epistemology. But we also need to take a second step. We need to accept that reality is multiple. Although this step is perhaps even more difficult, because of our common conceptual and ideological inertia, it solves a number of problems that have plagued eminent theorists who struggled with the problem of citation theory. If reality is multiple, we no longer need to search for one integrated theory that can explain the whole process of citing and citation. We can allow for a number of partly contradictory, and partly overlapping *sets* of citation theories, each emerging in a particular set of knowledge practices. The quest for a citation theory turns out to be the same type of chimera as the quest for universal transportable knowledge.

Let us delve into the details and reinterpret the process of citing and citation in these material semiotic terms. A scientist who cites a journal article or book written by another scientist creates by this very act a new object (different from the cited work). This object, the embedded reference, is tightly connected to the citing text. However, the scientist usually also creates another object as a consequence of his citing act: the reference in the bibliography. If she uses bibliographic software, this is done automatically. This second object is in principle still connected to the situated meaning of the first object but can be easily disconnected. This happens, for example, whenever another scientist copies the listed reference in his own bibliography or list of books to read. It is this second object which is the raw material for the citation indexer. In the indexing software, the reference is first inverted by sorting each citation according to the text that it cites, rather than by the citing texts from which it originated. This act of semiosis creates the third object, the individual citation (Wouters, 1999). Then, the citations are grouped together and counted. This results in the fourth citation object, the citation as part of the citation index. Note that the last two objects are not created by the citing scientist but by the citation indexing process. The citation indexes are sold commercially by the citation indexer or provided freely on the web as stand-alone products. They can also be sold or provided as part of more comprehensive research information systems. The citations thus provided are meaningful to the stakeholders involved, otherwise there would be no demand for them. But this meaning cannot be the situated meaning of the embedded reference since the citation objects have been created in a semiotic process that involves decontextualization. The original context has been lost. Therefore, the citation needs to be recontextualized in order to function as a social object. This happens in two different ways or modalities (see below) and it will no longer be a surprise that this means that actually two different types of citation objects are enacted in this process of recontextualization. And this increasing variety of citation objects all exist at the same time—they engage in

complex interactions with each other, as well as with other active objects such as references, evaluation committees, deans, and researchers engaged in knowledge creation.

4 Citation Link Object

In the first modality, the citation enters into a one-to-one relationship with the cited text (the right hand side of the Peircean triad): the object is the link between the cited text and the citing text. Let us call it the citation link object. In this recontextualization, the citation is literally the inversion of the reference as part of the bibliography. This object can play a variety of roles and take on a variety of shapes. It can be an instrument for the cited author to locate the exact place at which his work has been cited by the citing author. This is a form of inverse snowball sampling and Eugene Garfield saw it as the main use of the citation index (Garfield, 1955). In this form, the citation is a novel bibliographic tool that enables scientists to go forward in time in searching for literature, rather than only backwards in time (as is the case in traditional snowball sampling using the bibliographies). Interestingly, this means that the citation link object can be used to retrieve the situated meaning of the embedded reference from which the citation link object emerged. This is an important cause of the possibility for the citation link object to appear as simply another format of the reference. In other words, because of this potential as a search tool, which was actively advertised by Garfield in the early years of the citation index (Garfield, 1955), the process of decontextualization which creates the citation from the reference is made invisible. The citation can pose as actually being nothing else but the reference and this has given a very strong boost to the process of naturalization of the sign citation. It has made this process almost seamless.

This citation link object can also be the building block for creating maps of science. In this endeavor, all publications in a given area are presented as nodes in a network and their configuration is determined by the patterns of links between the nodes. These links can be any kind of object: the authors, the references, shared terminology, publication years, and, since the emergence of citation indexes, also the citation relations. The direct citation link is the most basic citation link object, but it can easily morph into much more complicated citation objects. For example, co-citation links are links between two documents that are each cited by a common third document, although they may not have other types of links between them, e.g., they may not cite each other (Small, 1973). The inverse of the co-citation link object is bibliographic coupling which is to the listed reference

in the bibliography what co-citation is to the single citation object (Kessler, 1963). Both direct citation links, co-citation links, as well as bibliographic coupling links are routinely used to create maps of science (Boyack, Klavans, & Börner, 2005; Small, 1977; Waltman, van Eck, & Noyons, 2010) (for more on science mapping, see Ginda, Scharnhorst and Börner, this volume).

5 Citation Number Objects

In the second modality, the citations to a cited text are aggregated and counted. The citation is now recontextualized not as the link between two texts, but as an attribute of the cited text. The cited text suddenly has a novel property: its citation frequency. The citation object is in other words recontextualized as metadata of the cited text. Moreover, it has now become a number and is therefore also recontextualized from the domain of textual strings and links into the domain of numbers and statistics. As a result, the possibilities of manipulation that the sign system of numbers has developed in the course of the last couple of centuries are now also available to the citation sign. Perhaps most importantly, the citation is thereby made commensurable and comparable. Cited texts can be compared in terms of their number of citations. These numbers can be added, leading to the number of citations of a particular oeuvre. As a result, the citation does not only become an attribute of the cited text, but also of the cited author. These aggregations can go on from author, to research group, institute, university, and country. At all levels of aggregation of the scientific system, we have a shadow reality of citations that hover over the actors (texts, knowledge creators, as well as institutions) as objects to which they can attach as well as detach. Practice, meaning, and context are intimately intertwined, they define each other, and it is therefore less fruitful to try to separate them into different domains of the symbolic representation and the real world. The reality of citation is indeed enacted within the double context of knowledge creation and research process evaluation.

6 Evaluative Bibliometrics

Evaluative bibliometrics as a field is the result of the emergence and enactment of this new citation reality as an added reflexive layer to the fabric of the scientific and scholarly system. Often, bibliometrics is seen as a data-driven field that has been "captured" by the creation of the Science Citation Index. This is a rather simple and poor explanation of the vastly richer semiotic reality that underlies

the field of evaluative bibliometrics. If we take semiotics seriously as an analytical lens, we have a much better framework to understand the historical development of the field and the pervasive influence of the citation index on the way quality is defined in the scientific and scholarly system.

To understand this role of the citation, we must pay close attention to the subtle but important difference between the two modalities of the citation object. The first form, that of the link between the cited and citing text from the position of the cited text, is still only two steps removed from the situated meaning of the embedded reference. It can be used to retrieve this situated meaning. It can also be used to try to find similar locations of situated meaning in other texts. This use of the citation sign is often firmly embedded within the context of knowledge creation, for example when a researcher wishes to find traces in the literature of researchers working on the same topic or thinking along the same lines. In the second form (the number that has become an attribute of the cited text), the citation object is resolutely disconnected from the situated meaning of the embedded reference which was at the origin of its semiosis. It is not the specific meaning of the citing author to which it refers. As an object, it exclusively represents the fact that a certain author or a certain number of authors has cited the text. What that means is not prescribed by the citation object itself. In this sense, we can claim that the citation included in the citation index as such is still an under-defined object or perhaps better a proto-object. It will only become a fully-fledged functional socio-material object if it materializes into a *specific* citation object, such as a specific indicator (e.g., the normalized number of citations per paper). In other words, the citation object in its second modality (as number) can be compared to a *stemcell object* which has as its main function of morphing into one of a large variety of concrete citation objects within specific citation, evaluation, and knowledge creating contexts.

This means that the development of specific citation theories that can explain the structure and role of citation objects in their various forms requires the inclusion of the specific practices and institutional contexts that are co-enacted. It clearly does not make sense to justify citation counts in an evaluative context by claiming that the number of citations is like a vote by the scientific community about the value of a particular piece of work or of a particular author. What the citation means cannot be extracted from itself, because the citation object as number is a decontextualized and underdefined proto-object. At the same time, it is also not the case that the meaning of the citation is completely determined by the context of evaluation. The meaning of the citation is not completely arbitrary and it does make a difference whether the number is higher or lower in a particular context. The values and shapes that the citation object can adopt are constrained by the patterns among the references from which the citation objects

emerge. Although the citation indexer is the producer of the citation objects, the citation indexer does not determine its values or shapes. In this sense, citation analysis is as objective as an analysis can be.

7 The Material Impact of Citation Objects

The system of citation has delivered key objects that are manipulated in this process of research evaluation and assessment. The development of the governance of research enabled and stimulated the uptake of these new objects (Wouters, 2014). The increased scale and specialization of scientific work as well as its increased role as a key instrument of production, reproduction, and distribution in current globalized capitalist economies have created a complex structure and introduced new dynamics into the management of science. The "social contract" that gave the scientific communities a relatively large amount of freedom (built on a combination of the notion of "academic freedom" in traditional academia and the promise of great economic and social progress—as well as profits—thanks to programmatic research) has been breaking down since the early 1970s. As a result, traditional forms of accountability and quality control in the form of various types of peer review have been supplemented as well as displaced by audits and accountability by science funders and stakeholders external to the scientific communities. The demand for more transparency in the organization of research than provided by traditional peer review forms has stimulated this development. These audits have created an increased demand for both quantitative and qualitative indicators as well as for forms of external expert review. The result is a complex dynamic of interactions between control by the relevant scientific communities and elites and control by various stakeholders. Who counts as a scientific expert or representative of a stakeholder community is not given a priori but is permanently redefined in these processes. Scientists play different roles in the interactions: as researchers, as experts, and as representatives of various stakeholder interests. How to combine these different roles and make them cohere is one of the challenges that researchers face in this complex science system.

National research systems have introduced a variety of performance based funding mechanisms, both based on peer review (such as the UK REF) and bibliometric evaluations (such as the Nordic bibliometric indicator) (Hicks, 2012). Currently, national systems vary in two dimensions: the tightness with which they couple performance assessment to funding and the degree to which they use metrics in addition to peer review (peer review based judgment is still the default). All combinations of these two dimensions occur. But even in systems without a

direct or tight coupling between funding and assessment outcomes, assessments have direct material consequences as the reputation of a research group or senior researcher is related to the outcome of the assessment. The systems only vary with respect to the details of this feedback loop and the time it takes for this cycle to complete. Countries with a relatively larger share of block funding may allow researchers more time to recover from a low rating in the regular assessment, provided that the block funding itself is not directly related to assessment outcomes. But in the long run, researchers will have to have high scores or leave the scientific system. This means that indicators do not only influence the judgment of past work, but they indirectly shape the possibilities for future research via this feedback loop. In other words, by influencing the research agenda of tomorrow they have direct material consequences for researchers and for science in general. Therefore, it seems insufficient to analyze citation indicators only symbolically without attention to materiality and economics. This again underlines that the material turn in semiotics may be fruitfully employed in scientometrics.

What does this mean for the position of citation theories? The field of scientometrics has developed a variety of citation theories, all of them interesting, none of them completely satisfactory or uncontested. One of the main lines of argument has been that citations reflect the referencing behavior of researchers (see above) and that therefore the citation rate of a paper, an author, a research group, an institute, and a country reflect the use of, and response to, the work by the relevant scientific community. This response can then subsequently be seen as a proxy for either quality or scientific impact (this varies in different citation theories). A supporting argument has been that at the level of the individual paper many factors other than the quality of the paper may influence the choice of references by the author, but that these factors cancel out at higher levels of aggregation (van Raan, 1998). In a recent paper, we have shown that this expectation may be a belief that is sometimes unsubstantiated (Waltman, van Eck, & Wouters, 2013), but it is still implicitly used by most citation analysts. The inquiry of motivations and factors influencing the references in scientific papers is still an active line of research in the area of citation theories (Bornmann & Daniel, 2008). From a semiotic perspective, however, this line of attack is misdirected because it assumes that references and citations are identical whereas from a semiotic perspective it is clear that this cannot be the case.

By including the whole reputational cycle in research in a material semiotic analysis, an empirical analysis based on this theoretical framework would be akin to Lenoir and Ross (1996). Lenoir and Ross (1996) aim to understand the way science functions as a disunified enterprise with the help of historically grounded semiotics: "We intend to demonstrate that the power of a sign, a representation, or an interconnected set of representations to support scientific work is not merely

a function of their own internal logic but also of their capacity to forge rhetorical links to representations in other domains by drawing upon metaphor as well as repertoires of tropes and narrative structures." Lenoir and Ross (1996) developed this to show how natural history museums are able to create artificially constructed sites to create a "meaningful nature". In the same vein, citation theory should analyze the main tropes and narrative structures in selected case studies of research evaluations (drawing on either documentary analysis or ethnographic field notes). Such an ethnographically grounded approach to the problem of citation theory would contribute to the work that aims to dismantle the artificial and unhelpful barrier between quantitative and qualitative work in scientometrics and science and technology studies (Wyatt et al., 2015).

Acknowledgment

Cited References

Berg, M., & Mol, A. (1998). *Differences in Medicine: Unraveling Practices, Techniques, and Bodies*. Durham and London: Duke University Press.

Bornmann, L., & Daniel, H. D. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, *64*(1), 45–80.

Boyack, K. W., Klavans, R., & Börner, K. (2005). Mapping the backbone of science. *Scientometrics*, *64*(3), 351–374.

Cronin, B. (1984). The citation process. The role and significance of citations in scientific communication. *London: Taylor Graham, 1984*, *1*.

Cronin, B. (2000). Semiotics and evaluative bibliometrics. *Journal of Documentation*, *56*(4), 440–453.

Dahler-Larsen, P. (2013). Constitutive Effects of Performance Indicators: Getting beyond unintended consequences. *Public Management Review*, *16*(7), 969–986.

Day, R. E. (2005). Poststructuralism and information studies. *Annual Review of Information Science and Technology*, *39*(1), 575–609.

Day, R. E. (2014). "The Data—It Is Me!"("Les données—c'est Moi!"). *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact*, 67.

Eco, U. (1976). *A theory of semiotics* (Vol. 217). Indiana University Press.

Egghe, L., & Rousseau, R. (1990). Introduction to informetrics: Quantitative methods in library, documentation and information science.

Garfield, E. (1955). Citation indexes for science. A new dimension in documentation through association of ideas. *International journal of epidemiology*, *35*(5), 1123–1127.

Gluck, M. (1997, August). Making sense of semiotics: privileging respondents in revealing contextual geographic syntactic and semantic codes. In *Proceedings of an international conference on Information seeking in context* (pp. 53–66). Taylor Graham Publishing.

Haraway, D. J. (1991). *Simians, cyborgs, and women: The reinvention of nature*. Routledge.

Hicks, D. (2012). Performance-based university research funding systems. *Research Policy*, *41*(2), 251–261.

Irvine, J. T. (1989). When talk isn't cheap: Language and political economy. *American ethnologist*, *16*(2), 248–267.

James, W. (1898). *Philosophical conceptions and practical results*. The University Press.

Keane, W. (2003). Semiotics and the social analysis of material things. *Language & Communication*, *23*(3), 409–425.

Kessler, M. M. (1963). Bibliographic coupling between scientific papers. *American documentation*, *14*(1), 10–25.

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Harvard University Press.

Latour, B. (1988). *The pasteurization of France*. Harvard University Press.

Latour, B., & Woolgar, S. (1986). *Laboratory life: The construction of scientific facts*. Princeton University Press.

Law, J. (2004). *After method: Mess in social science research*. Routledge.

Lenoir, T., & Ross, C. (1996). "The Naturalized History Museum." In *The Disunity of Science: Boundaries, Contexts, and Power*. Stanford: Stanford University Press: ,370–97.

Leydesdorff, L. (1998). Theories of citation?. *Scientometrics*, *43*(1), 5–25.

Luukkonen, T. (1997). Why has Latour's theory of citations been ignored by the bibliometric community? Discussion of sociological interpretations of citation analysis. *Scientometrics*, *38*(1), 27–37.

MacRoberts, M. H., & MacRoberts, B. R. (1996). Problems of citation analysis. *Scientometrics*, *36*(3), 435–444.

MacRoberts, M. H., & MacRoberts, B. R. (2009). Problems of citation analysis: A study of uncited and seldom-cited influences. *Journal of the American Society for Information Science and Technology*, *61*(1), 1–12.

Mol, A. (2002). *The body multiple: Ontology in medical practice*. Duke University Press.

Narin, F. (1976). *Evaluative bibliometrics: The use of publication and citation analysis in the evaluation of scientific activity* (pp. 206–219). Washington, D. C: Computer Horizons.

Nicolaisen, J. (2007). Citation analysis. *Annual review of information science and technology*, *41*(1), 609–641.

Nicolaisen, J., & Frandsen, T. F. (2008). The reference return ratio. *Journal of Informetrics*, *2*(2), 128–135.

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for information Science*, *24*(4), 265–269.

Small, H. G. (1977). A co-citation model of a scientific specialty: A longitudinal study of collagen research. *Social studies of science*, 139–166.

van Raan, A. F. (1998). In matters of quantitative studies of science the fault of theorists is offering too little and asking too much. *Scientometrics*, *43*(1), 129–139.

Walker, J. A. (2014). Comments on Umberto Eco's Book "A Theory of Semiotics". *Leonardo*, 317–319.

Waltman, L., & van Eck, N. J. (2012). Source normalized indicators of citation impact: An overview of different approaches and an empirical comparison. *Scientometrics*, *96*(3), 699–716.

Waltman, L., van Eck, N. J., & Noyons, E. C. (2010). A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics*, *4*(4), 629–635.

Waltman, L., van Eck, N. J., & Wouters, P. (2013). Counting publications and citations: Is more always better?. *Journal of Informetrics*, *7*(3), 635–641.

Wouters, P. (1998). The signs of science. *Scientometrics*, *41*(1), 225–241.

Wouters, P. (1999). The citation culture.

Wouters, P. (2014). The Citation: From Culture to Infrastructure. *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact*, 47.

Wyatt, S., Milojević, S., Lucio-Arias, D., Park, H. W., Šabanović, S., & Leydesdorff, L. (2015). Quantitative and Qualitative STS: The Intellectual and Practical Contributions of Scientometrics. *Available at SSRN 2588336.*