The Relevance of Citations

A case study of stratospheric ozone monitoring

JOAKIM PHILIPSON
A contextual citation analysis was performed on a sample of 68 articles citing either or both of two key papers from the research field of stratospheric ozone monitoring.

The theoretical approach of each article was assessed as either chemical or dynamical, or was left unspecified. Functions of citations to the two key papers were classified into 13 categories, seen to express various kinds of relevance relationships. Topical relatedness between cited and citing articles was studied in terms of number of shared descriptors and the union overlap measure. A certain theory-dependence, in the sense of a relatively high degree of correspondence between the theoretical approach of citing and cited articles, was observed in the material. The correlation between topical relatedness and citation relationships appeared to be weak.
Mångfaldigande och spridande av innehållet i denna uppsats – helt eller delvis – är förbjudet utan medgivande av författaren/författarna.
1 INTRODUCTION

1.1 PROBLEM AND HYPOTHESES

2 RELEVANCE

2.1 HISTORICAL OVERVIEW
2.2 TOPICALITY
2.3 NON-RELEVANCE

3 CITATION ANALYSIS

3.1 HISTORICAL BACKGROUND
3.2 CITATION FUNCTIONS
3.3 NON-CITEDNESS

4 RELEVANCE AND CITATION ANALYSIS

4.1 THE RELEVANCE OF CITATIONS
4.2 METHODOLOGY

5 A CASE STUDY: MONITORING OF THE OZONE LAYER

5.1 BACKGROUND: STRATOSPHERIC OZONE MONITORING
5.2 THE MATERIAL
5.3 CLASSIFICATION OF CITATION FUNCTIONS
5.4 RESULTS

6 DISCUSSION

7 SUMMARY AND CONCLUSIONS

REFERENCES
Acknowledgements

There are three persons in particular that I would like to thank for helping me in writing this thesis:
First, Jan Nolin at the Department of Theory of Science, Göteborg University, who not only provided necessary background knowledge about stratospheric ozone monitoring through his dissertation, but also was so kind to let me browse through his files and copy valuable material.
Secondly, Jakub Srebro, chemical engineer at the Apoteksbolaget in Stockholm, who answered some of my naive questions about the properties of esters.
Last, but not least, Ann Berman, Stockholm, who in no time produced an expert translation from Spanish into English of some key passages of an important document. Their contribution is fully and gratefully acknowledged.

Valuable comments on earlier versions of the manuscript were given by Anders Malmsjö, University College of Skövde, to whom I am also indebted.

The no doubt numerous errors, linguistic as well as logical, that I have failed to remove from the text, are all of my own produce.

Borås, October 1996
Joakim Philipson
1 Introduction

Relevance is one of the most fundamental concepts in information science. The sheer volume of literature on the subject is telling evidence for the centrality of the notion. In his early review article Saracevic (1976) states as a fact:

Relevance was the key notion in the emergence of information science, and it has remained the basic notion in most of its theoretical constructs and all of its practice.

(op.cit., p. 85)

However, despite the large amount of research that has been devoted to clarification and explication of the concept, there still seems to be lacking a consensus among information scientists as to the exact nature of relevance, as Froehlich (1994, p. 24f) noted more recently.

This study does not aim at proposing yet another definition of relevance, the proliferation of definitions already being daunting. Rather, it constitutes an attempt at assessing the possible relevance of citations appearing in scientific papers within a particular field of research, the assumption being that such citations can often be seen as a kind of recorded relevance judgments, made by experts.

One of the main sources of inspiration for this study has been two recent articles by Rebecca Green and Carol A. Bean (Green 1995a, 1995b) in which topical relevance relationships are analyzed in detail and a preliminary typology is presented. The articles argue convincingly for a more complex notion of topical relevance than the traditional view based on "topic matching", and they seem to suggest a more profound impact of theoretical, intra-scientific considerations on relevance judgments than has generally been assumed by information scientists and librarians. Green (1995a) also points explicitly to citation analysis as a method of unearthing relevance relationships that are not based on "topic matching".

In the remainder of chapter one the research problem and the hypotheses that seem to emerge from the case-study are presented. In chapter two an overview of previous research on relevance is given, followed by a brief discussion of two concepts closely connected with relevance, viz. topicality and non-relevance. The third chapter deals with citation analysis in a similar vein, giving a brief historical overview, followed by a discussion of citation functions and of some of the reasons for not citing. Chapter four analyzes more closely the relationship between relevance and citations, after which the methodology used for the empirical study in the following chapter is introduced. In chapter five the results of the case-study, a qualitative citation analysis of a pair of articles from two scientific journals, are presented. The case-study is from the field of research concerned with monitoring of the stratospheric ozone layer. It is to a large extent inspired by a recent dissertation, presented at the Department of Theory of Science, Göteborg University (Nolin 1995), which provided much useful background information. In chapter six, the results of the case-study are discussed and put into
perspective. Chapter seven, finally, contains a summary and some possible conclusions that could be drawn from the results of the case-study.

1.1 Problem and hypotheses

Despite its centrality in information science, the concept of relevance is still lacking a generally accepted, clear-cut definition. Part of the confusion surrounding the notion may stem from an insufficient insight on the part of information scientists into the nature of scientific knowledge and the process during which it is produced. Naive empiricism, the view of science as a more or less linear process of cumulation of knowledge, as more and more facts are discovered, has long since been abandoned by the philosophers of science, and has been replaced by more sophisticated and complex views. However, in much of what has been written on the concept of relevance in information science, there still seems to be implicit some kind of empiricist view. In particular, the still common view in information science, that information may be relevant to a topic or to a problem often seems to entail an empiricist view of knowledge, according to which a scientist begins by collecting all the relevant information or "facts" about his or her object of study, and only then sits down to theorize about it.

In contrast with this view, a hypothesis suggested by this study is that relevance is generally "theory-dependent". No scientist begins his or her research from scratch. Rather, she or he most often starts with some kind of hypothesis or theory about the nature of the object of study, and this hypothesis or theory also entails some notion about what kind of information would be potentially relevant to his or her research. As Hempel (1966) has observed, what kind of information it is rational for a scientist to gather in his or her research is determined by a conjecture or hypothesis about a tentative answer to the research question at hand, rather than by the research problem all by itself. Stating the point somewhat differently: the problem all by itself does not tell me where to look for an answer. To direct my search, I must have at least some idea about what a potential solution to my problem should contain. This ought to mean,

1 Cf. e.g. the works of Popper, Lakatos, Kuhn, Feyerabend, Toulmin et al. More recent names that have contributed substantially to our understanding of how science works are e.g. Bloor, Barnes, Knorr-Cetina, Collins, Latour, Chubin and Restivo.
2 Hempel is still very much associated with an empiricist view of science, albeit a more sophisticated version of it.
3 There seems to be at least an implicit awareness of this with some researchers in library- and information science, who stress the importance during a reference interview of asking the patron about the intended use of the information sought. Some researchers also explicitly recommend asking for a description of an acceptable potential solution to the problem at hand. However, their interest seem to be more to find out about the formal characteristics of such a solution, in which type of source the answer could be expected to be found etc., and not so much about the content of a hypothetical answer (cf. e.g. Kuhlthau 1993).
that the relevance of a given piece of information to a research problem can only be fully assessed with respect to some hypothetical answer to the problem at hand. Thus, if the hypothesis put forward in this thesis can be corroborated, the librarian or information specialist, rather than asking the scientist only about her or his subject or field of research, should perhaps also ask: What is your theory?

However, in order not to beg the question, and since the reasons motivating both the search for and the use of information are manifold, henceforth the term "purpose" will generally be used instead of either "question", "problem" or "hypothesis". Thus, for example, the expression "relevant to a purpose" will be employed instead of "relevance to a problem". The advantage of this mode of expression, although it may not conform completely to the standard English usage⁴, is that it appears to be more general. Obviously, the purpose for which a scientist engages in research, writes a scientific paper and cites other documents may range over a multitude of possibilities. She or he may indeed want to solve a problem, or seek confirmation for her or his favorite hypothesis (or, alternatively, seek to refute her or his adversary's favorite hypothesis). But it may also be that he or she simply wants to impress a certain professor, get a grant from a certain body of decision-makers, or that he or she is doing it "just for the fun of it".

By "theory" or "theoretical approach" in this thesis is meant, if nothing else is stated, a set of beliefs shared by a group of scientists within a research field. "Theory", in this sense, corresponds roughly to the second component of a "disciplinary matrix", which Kuhn (1970) termed a "metaphysical paradigm", providing as examples such beliefs as: "heat is the kinetic energy of the constituent part of bodies" and "the molecules of a gas behave like tiny elastic billiard balls in random motion" (op. cit., p. 184). It is important to remember also in this context, that a theory or a metaphysical "paradigm governs ... not a subject matter but rather a group of practitioners" (op.cit., p. 180). Thus, as will be seen in the case-study performed here, it is possible to have different, competing theories within one and the same field of research.

When applying the concept of theory in the case-study, however, the analysis focussed exclusively on the actual documents studied, rather than on their authors, the scientists working in the field of research studied. That is, it was assumed that the "theory" or "theoretical approach" in the sense given these terms above, could be effectively "extracted" by an analysis of the text of the articles appearing in the case-study, in view of the background knowledge of the field of research under consideration that was provided by Nolin (1995). This means, that the decision to classify an article as representing the one or the other theoretical approach was based solely on the contents of that article, disregarding also previous or later known works by the same author(s). In order to avoid forcing a theoretical approach onto articles that proved difficult to classify in this way, an utterly cautious attitude was adopted, so that more than a third

---

⁴ According to The Advanced Learner's Dictionary of Current English (2nd ed., Oxford University Press 1963) the expression "to the purpose" means "useful for one's purpose; relevant".
of the articles analyzed were actually left "unspecified" in this respect. The specific grounds for the classification of articles as for theoretical approach are presented below in section 5.1.

An alternative strategy for getting at the beliefs of scientists would, of course, have been to ask them simply, that is, making a survey or performing interviews with authors of scientific papers. The reasons why such a strategy was not adopted for the present study are stated in section 4.2, dealing specifically with methodology.

The beliefs that make up a theory tell scientists, among other things, what kind of answer they could expect to a research question, and where to direct the search. Hempel (1966) provides the example of Semmelweiss, who found the cause of death in childbed fever only by assuming first as an hypothesis that it had something to do with contamination from corpses. Only then could he proceed to gather the information judged relevant with respect to the hypothesis, by investigating what happened when medical doctors and students disinfected their hands before examining their patients.

Now, it is clearly impossible to prove that every relevance judgment made everywhere at any time is influenced by theory, in the sense used here. The material analyzed in the case-study is too small even for allowing any general inductively inferred conclusions. Although care was taken in the selection of articles from which the citations were drawn to get a reasonable distribution over the time period studied, with respect to the development within the field of research under consideration, there is no claim as for their possible representativity.

What this thesis sets out to do is the more modest task of showing or making probable

(1) that there are at least some cases, in which topic matching relationships are not necessary for explaining relevance judgments, as manifested by citations to articles in scientific journals, and

(2) that theoretical approach does indeed appear to influence relevance judgments, again as manifested by citations to scientific papers, i.e. that there is an apparent correlation between theoretical approach of citing and cited documents

To accomplish this it must also be shown that the citations analyzed can in fact be seen as based on relevance judgments, rather than being instances of citations "honoris causa" or having other non-relevance based explanations.5

5 Note, however, that the cases referred to in (1) and (2) respectively are not necessarily identical. In the study performed here, there is at least one case of "misfit" between citation relationship and topic matching, which is also an apparent "misfit" between citation relationship and theoretical approach.
In order to put the hypotheses formulated in (1) and (2) to a real test much further research, and on a more extensive material is required. This thesis contends itself with exemplifying and illustrating them by means of a case-study from a seemingly narrow, although clearly cross-disciplinary field of research; that of stratospheric ozone monitoring.

The method used for the case-study will be expounded in detail in chapter four, in connection with the theoretical discussion of the relationship between relevance judgments and citation practice. In short, the method involves three steps:

(i). Select from scientific journals a pair of articles, representing different theoretical approaches, but between which a topic matching relationship is holding; the criterion for topic matching being specified by the requirement of having at least two descriptors in common, as assigned to the articles in at least one of the databases in Knight-Riddler's DIALOG information retrieval service. Further, the selected articles should be as close as possible in time and location. This is in order to minimize the influence on citation behavior of factors such as differences in exposure, impact and prestige of the journals in which the articles appear.

(ii). Find out, by means of the SciSearch database (Dialog file no. 434), which documents have cited the selected articles and obtain a fair portion of them (in case there are too many citing documents, a selection is made by limiting the search to specific publication years).

(iii). Analyze the obtained documents in order to determine
(a) functions of the citations to the initial pair of articles in each case, excluding from the next step in the analysis suspected instances of 'honorary' and other non-relevance based citations.
(b) theoretical approach of citing articles, and how it matches with that of the cited source.
(c) topical relatedness, both in absolute terms, in accordance with the criterion for topic matching specified in (i) above, and as expressed by the union overlap measure (explained below section 4.2), between citing article and the initial pair of articles.

The overall purpose of this thesis, then, is to assess the possible relevance relationships between citing and cited documents within a particular field of research during a limited period of time, with the special objective of making plausible the hypothesis that relevance judgments in science also involve important theoretical considerations. More specifically, by choosing a pair of scientific articles that are close both in topic and time, but that show significant differences as to their theoretical approach and rate of citedness, the idea was to uncover, by means of contextual citation analysis, relevance relationships the explanation of which do not require matching topics of citing and cited documents.

6 Ideally, they should be published in the same issue of a scientific journal, to make sure that a potential citer is likely to have been exposed to both articles. In the present case-study, however, it was not possible to fulfill this requirement, although the two articles selected are "close enough".
Documents containing relevant information have been characterized as "moving targets whose dominant components communicate to us and affect our behavior" (Ottaviani 1994, p. 264). While recognizing the changing, volatile nature of relevance, this thesis sets out to study two such "targets", which at one point in time seemed closely related, and how they have moved about, been viewed, aimed at and hit from different angles at different points in time.
2 Relevance

Relevance has been a central concept of information science already from the inception of the field as a separate academic discipline. This is not difficult to understand. Information science is concerned with the design and evaluation of information systems. For both of these tasks relevance is evidently a key notion. The raison d'être of an information system is to satisfy the information needs of its users, that is to provide them with relevant information. Intuitively, then, relevance can be seen as some kind of relation holding between an information source and a user need. Or, as one of the numerous proposed definitions of the concept has it, relevance is "a measure of the effectiveness of the contact between a source and a destination in a communication process" (Saracevic 1976, p. 91).

Thus, on a very primitive level, relevance could be described schematically by the following figure:

![Figure 1](image)

Two basic elements are involved, an information source S and an information user U, with the arrows between them representing the communication process by which the information reaches the user. There are two arrows, one in each direction, to indicate that in order for a relevance relationship to obtain it is necessary not only that the particular piece of information reaches the user, but also that the user actively selects this particular piece of information from others by judging it relevant to his or her information need. The scheme is a very crude one, since for example the information system that mediates between the source and the user has no representation in it. For the purpose of this study, however, it is sufficient. What will be dealt with here is, precisely, the relationship holding between the user of information and the information sources used.

Different schools of thought have leaned more or less heavily against either of the two poles in Figure 1. It is probably safe to say that the general trend in the development of the thinking on relevance in information science during the last fifty years has been a gradual shift of focus from the source to the user. The following overview of the historical development follows closely in the footsteps of Saracevic (1976) and Froehlich (1994). It is not intended to be a complete account of what actually happened in the field during the last fifty years, but rather, to provide a background for the theoretical discussion that will follow later.
2.1 Historical overview

Saracevic (1976, p.126ff) discerns at least seven major different views of relevance, some appearing simultaneously, some succeeding each other in the historical development of information science. First to appear on the scene, after the Second World War, was (1) the so called "system's view", according to which relevance was largely a question of manipulating an information system, of file organization and searching strategies (op. cit., p.96). The principal aim was to reduce the number of "false drops", that is non-relevant, unwanted retrieval, by refinement of the information system through better coding, classification and indexing. The user was largely left out of this discussion. The concept of relevance employed was more or less intuitive and was actually presupposed to be known in the definitions of the retrieval measures precision and recall7, that were developed in the mid fifties.

The uneasiness that was beginning to be felt among information scientists with the "system's view" came to expression at the International Conference on Scientific Information in 1958. Hillman (1964), later summing up some of the most critical points raised at the conference, concluded:

We are not so much interested in discovering what items are defined by a system to be relevant to a query, but how the concept of relevance is first to be defined, and only later how theories can be constructed in accordance with this definition to generate those systems capable of retrieving all items relevant in this logically prior sense.

(op.cit., p. 27)

Thus, the "system's view" had begun in the wrong end. Rather, to arrive at a plausible concept of relevance research had to start by asking for the grounds on which relevance judgments are made, independently of the system. Other insights emerging after the 1958 conference concerned the multivalued, changing nature of relevance (cf. Saracevic 1976, p.97). These developments seem particularly to have gained momentum with the methodological debate that followed in the wake of the Cranfield Project in the late fifties and early sixties. The aim of the project was still very much in line with the "system's view", large-scale testing of IR-systems, but the resulting debate concerned precisely one of the most critical points of this view: if performance of a system was to be measured in terms of relevance, then relevance must be determined independently of the system. But how, then, should relevance judgment be passed? And who should be the judges? As Saracevic (1976) notes, "the debate imperceptibly but completely shifted the problem of relevance from the system to the destination" (op.cit., p. 98). Thus, (2) the "destination's view" emerged and with it the user entered the scene.

7 Recall was defined as "the ratio of relevant answers retrieved to the total number of relevant answers in the file" and precision as "the ratio of relevant answers retrieved to the total number of answers retrieved" (Saracevic 1976, p.97)
In the midsixties, spurred by a 1964 conference that had been called together by the National Science Foundation, major experiments were performed\(^8\), aiming to assess the effects on relevance judgments of different user characteristics. The results seemed to point in a certain direction. It was found, perhaps not surprisingly, that among the most important factors that influence relevance judgments are users' subject knowledge and intended use of documents (cf. Saracevic 1976, p. 133f). All in all, about forty different variables affecting relevance judgments were identified as a result of these experiments (Schamber 1991, p. 239). Variables were subdivided into five major classes:

1. Documents and document representation
2. Queries, or "Information requirement statements"
3. Judgmental situations and conditions
4. Modes of expression [i.e. instruments by which relevance judgments are expressed, such as rating scales]
5. People (judges)

(Cuadra & Katter 1967, quoted from Saracevic 1976, p. 100)

The fact that documents and document representations were classified together clearly indicates the shift of focus that had occurred in information scientists' view of relevance. In the system's view it was precisely the relationship between documents and their representations in the system that was believed to affect relevance the most. By contrast, in the above classification scheme, there is a clear dominance of factors relating to users; four out of five classes can be said to concern, directly or indirectly, characteristics associated with the user, or with the judges. Relevance was more or less equated with relevance judgment. Psychology became an integral part of information science. This was reflected in numerous definitions of relevance that were produced during the sixties. According to Saracevic (1976) they very often conformed to the general pattern "relevance is the A of a B existing between a C and a D as determined by an E" (op.cit., p. 99f), where the three latter terms corresponded roughly to classes 1, 2 and 5 respectively in the classification scheme above, whereas A and B expressed the relation holding between them. A model example of this kind of definition would sound, then, "relevance is the measure of correspondence existing between a document and a query as judged by a person". The last word in all these definitions of relevance is thus given to the person passing judgment on the matter. Whereas the "system's view" had made relevance depend on the information system, the "destination's view" made it entirely dependent on the relevance judges. In both views, relevance became something relative, eluding an "objective", general definition.

However, parallel to this development, there were also attempts at closing in on relevance by other, more "objective" measures. One of them was bibliometrics and citation analysis, which will be discussed in chapter three. True, just like in the "destinations view", in citation analysis too, the user is still very much in focus, although in the latter case it is not his or her internal state of mind that is of interest, but

---

\(^8\) Two of the more prominent experimental studies were those performed by Rees & Schultz (1967) and Cuadra & Katter (1967). For further references to reports on these experiments, cf. Saracevic (1976, pp. 129-135).
rather the concrete manifestations, in the form of citations, of the relevance judgments she or he might pass. Bibliometrics, in fact, developed earlier than both the system's view and the destination's view, but was not directly concerned with relevance originally, although Bradford, who formulated one of the fundamental bibliometrical laws, already in the 1930's and 1940's was talking about articles "relevant to a subject" (Saracevic 1976, p. 89). Bibliometrical methods served mainly as tools of evaluation and description, be it of library collections, of scientific disciplines or even of individual scientists. The development of citation indexes, though, had something in common with the "system's view", in the sense that it sprang from an original endeavor to improve the performance of information systems (cf. Garfield 1979, p. 6ff). Some of the important results of bibliometrics, such as the laws governing the distribution of scientific literature, discovered by Bradford, Lotka, Zipf and others, were synthesized and interpreted in terms of relevance by Saracevic in an earlier work (1970). These "relevance related distributions" formed the basis of what he called (3) the "subject literature view" of relevance (Saracevic 1976, p. 105), according to which relevance was seen as a relationship between a subject, or a topical question on the subject, and the literature representing the subject (op.cit., p. 126).

Another of the more "objective" approaches towards relevance was (4) the formal or "logical view", which concerned itself mainly with the nature of the inference from search queries to documents retrieved, or rather, to the information contained in those documents. The inference relationship could be seen as either deductive or inductive, probabilistic. Cooper (1971) developed the former view; placing three formal restrictions on the search query (viz. that it be a yes-no, true-false question), the stored data (that it be expressed in logically well-formed sentences) and the retrieval system (that it be inferential, deducing direct answers to input questions) respectively, he defined:

A stored sentence is logically relevant to (a representation of) an information need if and only if it is a member of some minimal premiss set of stored sentences for some component statement of that need.

(op.cit., p. 24)

According to this definition, then, relevance to a search question is tantamount to membership in a minimal set of premisses from which an answer to the question can be deduced. The requirement that the set be minimal is necessary to avoid redundancy.

Wilson (1973), building on Cooper but preferring inductive logic to deductive defined what he called "evidential relevance" thusly:

an item of information $I_j$ is relevant to a conclusion $h$ in relation to premisses $e$ if the degree of confirmation, or probability, of $h$ on evidence $e$ and $I_j$ is greater or less than the degree of confirmation, or probability on $e$ alone. (op. cit., p. 460)

Evidential relevance, then, simply stated means that an information item either "strengthens or weakens a case", whereas "irrelevant information is that which does neither" (ibid.). This seems to accord well with Saracevic' view of relevance as a
measure of the change taking place in the destination file as a result of a communication of knowledge from the source to the user:

A source and a destination involved in the communication of knowledge both have, as one of the elements, a file (or files) where subject knowledge and /or its representations are stored in an organized manner (structured, associated, etc. Examples of such files are: memory, library collection, catalog, computer file, data bank, and store of sentences. Communication of knowledge is effective when and if information transmitted from one file results in changes in another. Relevance is the measure of these changes. (Saracevic 1976, pp. 92f; emphasis in the original).

Clearly, if an item is evidentially relevant in Wilson's sense, the user (e.g. the writer of a scientific paper) will either want to include, typically by means of a citation, the information provided by the item in his or her "file" (e.g. a journal article) to strengthen her or his case, or will likely have to remove or modify his or her hypothesis, if the information provided is seen, on the contrary, to weaken the case, that is, has negative relevance.9

Common to both the "deductive inference view" and the "probabilistic inference view" is their primary interest in the very nature of relevance relations; they are not so much concerned with attempts made by others to enumerate all the elements that enter into relevance relationships (Saracevic 1976, pp. 123f).

The split between "subjective", conditional views of relevance on the one hand and more objective approaches on the other was also making itself felt in the proposed "division of labor" between relevance and pertinence. This distinction entered information science with the mounting influence of psychology in the sixties, and was a result of the introduction of another distinction, that between a question and an information need (op. cit., p. 107). Goffman and Newill (1966) explained the relation between the concepts thus:

Relevance is the property which assigns certain members of a file (e.g., documents) to the question; pertinence is the property which assigns them to the information need. (quoted from Saracevic 1976, p. 108)

From this distinction already it can be seen that there may well be relevant answers that are not pertinent. Presumably, the individual information user is primarily interested in pertinent answers, that satisfy her or his need, whatever it may be. But an information system generally cannot look into the user's mind, it can only answer real questions. That is, at best, it can only offer relevant answers.

9 The author of a scientific paper could, of course, choose to ignore negatively relevant information altogether, thus leaving the file unchanged, but only at the risk of standing up against the harsh verdict of the scientific community. True, this risk may sometimes be only theoretical. Due to the information explosion in science, it might be impractical to control every instance of suspectedly "overlooked" or "ignored" counterevidence to a case.
Foskett (1972) later restated the distinction as one in which "relevant" should be taken to mean "belonging to the field/subject/universe of discourse delimited by the terms of the request, as established by the consensus of workers in that field; while "pertinence" should be taken to mean "adding new information to the store already in the mind of the user, which is useful to him in the work that prompted the request" (op. cit., p. 77).

Thus, the distinction between relevance and pertinence was transformed into a distinction between two kinds of knowledge, public and private, resulting in two different views of relevance. The first was (5) the "subject knowledge view", which "considered the relation between the knowledge on or about the subject and a topic (question) on or about the subject" (Saracevic 1976, p. 126). This view resembles closely the "subject literature view", (3) above, but Saracevic is careful to point out that they are not the same (op. cit., p. 111), the reason obviously being that not all existing knowledge is recorded in the literature. It may also be that some recorded knowledge has fallen into oblivion and is not known by the scientific community. Further Saracevic claims that "the subject knowledge view of relevance is fundamental to all other views on relevance, because subject knowledge is fundamental to communication of knowledge" (ibid., emphasis by Saracevic). This is easy to agree with. The accumulated subject knowledge within a scientific discipline or field of research obviously is one of the important factors determining what kind of problems to address, and, hence, what type of information will be needed to solve those problems. For example, the subject knowledge of a scientist that is about to write an article most likely affects the selection of references to other works used for the article.

The second view of relevance resulting from the distinction above, the one associated with private knowledge, was (6) the "pertinence or destination's knowledge view", which "considered the relation between the stock of knowledge at hand of a knower and a. subject knowledge; or b. subject literature" (Saracevic 1976, p. 127). This view is close to the "destination's view" in that it focuses on the information user, or rather his or her stock of knowledge. Saracevic actually views the destination's view as incorporating the "pertinence view", since the stock of knowledge at hand is but one of the factors possibly affecting the user's relevance judgment, whereas he values the latter view as being "much more to the point" (op. cit., p. 112).

The proliferation of definitions of relevance during the sixties, and the notorious vagueness still inherent in them, spurred also a growing skepticism and a disenchantment with the notion (ibid.). A number of alternative terms to replace or supplement "relevance" were suggested; terms such as "usefulness", "appropriateness" or "utility". The latter term, notably, was employed by Cooper, who distinguished carefully between "logical relevance" (or "topicality") on the one hand, and "utility", involving also such factors as the importance, credibility and intelligibility which the

---

10 The long-dormant work of Gregor Mendel in genetics is a case in point.
Logical relevance is only one of the factors affecting utility, although it is, according to Cooper, the single most important one (ibid.).

Cooper's ideas on utility were developed by others to form yet another view of relevance, (7) the "pragmatic view" (Saracevic 1976, p. 115). Although Cooper himself apparently preferred to restrict the use of the term "relevance" to stand for a logical relationship, the general trend in this "pragmatic thinking" on relevance in the early seventies seems to have been a certain relativization of the concept, permitting different kinds of relevances, involving different relations. Wilson (1973), notably, distinguished between logical or evidential relevance on the one hand and psychological relevance on the other. He then introduced the concept of "situational" relevance, in which the information user's beliefs (stock of knowledge) concerns and preferences over a range of alternatives was coupled with logical or evidential relevance. In the same line of thought Kochen (1974) distinguished between relevance as a relation between propositions and relevance judgments made by a user. Such judgments, according to Kochen, can be based on a notion of utility, which he defined formally by means of four axioms, all demanding that the user be capable of assigning an order of preferences to the documents in a collection or specifying the conditions under which he or she will remain indifferent in the choice between two documents (cf. Saracevic 1976, p. 116f).

Disregarding the practical difficulties that may result from a definition based on preferences, the problem with this type of notion is very much the same as that which affects the concept of recall (cf. above and note 3): it seems to require a more or less complete overview of and acquaintance with the collection of documents from which the selection is made. But obviously in most instances of information retrieval the conditions are not so favorable as to permit that. Rather, we often select information ad hoc, as we go along, while working on our different projects. Kochen does seem to recognize, however, that a change of preferences may occur with learning or with other variables changing (ibid.).

Summing up thus far we discern with Saracevic (op. cit., p. 119) three views of relevance that are clearly focussing on the right hand element in Figure 1, that is on the information user: (2) the "destinations view", concentrating on judgment, (6) the "pertinence view", concentrating on the user's stock of knowledge at hand, and (7) the "pragmatic view", concerned with the utility of the information to a user and his or her preferences. The four remaining views presented briefly above, viz. (1) the "system's view", (3) the "subject literature view", (4) the "logical view", and (5) the "subject knowledge view", appear to be more "objective", closer to the information source, and should consequently be placed somewhere in between the two poles in Figure 1, although it is hardly possible to determine their exact position or even the exact order between them.

It must be made clear already at this stage that elements from all these views of relevance are involved in citation analysis. For example, the pertinence view (6) has a point in the fact, that the stock of knowledge at hand of a scientist about to write an article for a scientific journal most likely affects the selection of references for the user attaches to the information retrieved, on the other (cf. Cooper 1971, pp.35f).
article. On the other hand, as will be seen in the case-study, the function of a citation does not seldom appear to be that of providing evidence in favor of the hypothesis put forward by the author of a scientific paper, something that seems to speak for the logical view (4) of relevance. And of course, exactly what pieces of information finally enter a scientific journal article in the form of citations is highly dependent on the judgment of the author, as the destinations view (2) would emphasize.

After the lively debate in the beginning of the seventies interest in the question of relevance seems to have faded for some time, as noted by Eisenberg (1988). A possible explanation is the above mentioned general disenchantment with the concept as such and the failure to arrive at a generally accepted, "final" definition. It is not until the late 1980's and the beginning of the 90's that the issue reemerges at the agenda again. Today, however, it appears to be a widely accepted opinion that relevance is a multidimensional concept, and that no single view can lay claim to having defined its true nature (Eisenberg 1988, Froehlich 1991). At the same time, the general tendency towards a user-centered notion of relevance has become even more marked, as can be seen e.g. from the articles in the special topic issue on relevance of JASIS, Journal of the American Society for Information Science, vol. 45 (1994): no. 3. Several researchers have conducted extensive surveys of user-centered criteria that affect relevance judgment (cf. e.g. Barry 1994, Howard 1994 in the same issue of JASIS). Hand-in-hand with this continued movement towards a user-centered notion seems to follow a growing awareness of the inadequacy of topicality as the basis of relevance judgments. However, it is sometimes not entirely clear what is meant by this term, although it has long been at the very heart of the debate over relevance. For this reason, and in order to prepare the ground for the case study, the next section will be devoted entirely to the concept of topicality.

2.2 Topicality

Topical relevance has often been associated or even equated with the "system's view" (cf. Froehlich 1994, p.125). Hence, the main critique directed against topical relevance has been that it ignores the user and the specific context in which relevance judgments are passed. Park (1994) summarizes some of the most common objections:

Topical relevance is context-free and is based on fixed assumptions about the relationship between a topic of a document and a search question, ignoring an individual's context and state of needs. It is a unidimensional view of users' information problems, disregarding the changing nature of the individual's information problem and its subsequent impact on the search. It fails to focus on the complexity of the the individual's background and task situation.(p. 136)

---

11 The multidimensional character of relevance was evident already in two studies from the sixties by Rees and Schultz (1967) and Cuadra and Katter (1967), which both discerned about 40 variables involved in relevance judgments. See Eisenberg (1988).
A similar critique has been formulated by Schamber (1991):

Only one definition of relevance does not relate directly to the user: the definition involving topicality, or the mechanical matching of terms describing topics of documents with topics of information requests. (p. 239)

In this last quotation can be seen also the common identification of topical relevance with topic matching. The most important contribution of Green (1995a,b) is perhaps to have shown that there are also other types of topical relevance than topic matching, viz. hierarchical and structural (or syntagmatic) relationships.

Hierarchical relationships are either taxonomic in character, that is "the topic of the user need and the topic of the cited passage are related as class and subclass, or... as class and class-member" (Green 1995b, p. 659). This latter kind of type-token relationship is often expressed by the provision of an example or instance of the class referred to. Another kind of hierarchical relationship, less common in the special case studied by Green, is partonomic (part-whole) in nature, typically corresponding to a user need of the form "What does X consist of?" (ibid.).

Structural (or syntagmatic) relationships are those where
the topic of the cited passage corresponds to a component within a conceptual syntagmatic structure (or occasionally to the structure at large),
while the topic of the user need corresponds to another component within the structure, or again, the structure at large (op. cit., p. 660).

A syntagmatic structure is an integrated conceptual unit, the components of which can be seen as "slots in a framework" (ibid.). The example provided by Green is that of a path, which has three components (slots): a point of origin, a point of destination and a link between them. A more general, and for the purpose of this thesis more relevant example would be that of citations serving as evidence in support of an hypothesis or a conclusion of a scientific paper. Such evidence, according to the "logical view" of relevance (cf. Cooper 1971; see previous section), can be seen as premisses in an inference (inductive or deductive) leading to the hypothesis or conclusion in question. The relationship between these premisses seem to be rather well described as being structural in nature, in the sense defined by Green (1995b). Consider the following extract from an article published in a botanical journal as an illustration:

Good estimates of the present stratospheric distribution of ozone and subsequent UV radiation are known (Koller, 1952; Dütsch, 1969; Cutchis, 1974). The total amount of ozone in the northern hemisphere is maximal in spring and minimal in fall. ... It is suggested that among flowering plants of the northern hemisphere, many of which have white or yellow flowers (Table 2), there has been convergent evolution in floral UV absorption. Yellow and white flowers are high in flavonoid pigments which strongly absorb UV light. The seasonality of UV radiation may be one major selective pressure. Yellow and white flowers comprise as much as 85% of an arctic flora (Kevan, 1972) ...

(Utech 1975, pp.26f)
Not being in a position to enumerate all the premisses involved in the inference leading to what appears to be the hypothesis put forward here, i.e. the third sentence in the just quoted passage, we may nevertheless discern some of the more important of these premisses. First, there is the observation of the seasonal variation of stratospheric ozone and the subsequent seasonal variation of ultraviolet radiation reaching the earth, leading to a spring maximum of stratospheric ozone and a subsequent spring minimum of UV-radiation in the northern hemisphere (since stratospheric ozone absorbs UV-radiation). Then there is the knowledge that yellow and white flowers are strong absorbants of UV-radiation. Finally there is the evidence of the predominance of yellow and white flowers in the northern hemisphere. Together these premisses make probable the hypothesis that UV-absorption ability has acted as a selective evolutionary mechanism for flowers in the northern hemisphere. It is important to note here that the different premisses come from different topical areas. The first three cited sources in the extract are entitled "Ultraviolet radiation", "Climate of the free atmosphere" and "Stratospheric ozone depletion and solar ultraviolet radiation on earth", apparently belonging to fields such as geophysics or climatology, whereas (Table 2) and (Kevan, 1972) refer to documents from the subject area of botany, with titles such as "Coloured illustrations of herbaceous plants of Japan (Sympetalae)" and "Floral colors in the high arctic with reference to insect-flower relations and pollination". (Note also that UV-radiation, which for the citing document evidently has clear implications for botany, does not seem to be a major concern of these latter documents, at least not judging from their titles.) Nevertheless, despite the differences in subject, the premisses obviously "fit" together, as slots in a framework. One describes certain environmental conditions. Another describes an important property of the object being studied, influencing its adaptation to the conditions described by the first. The third premiss, finally, describes the frequency of occurrence of the object being studied, thereby corroborating the importance of the property described by the second premiss. Together they make up an evidential structure, that would seem to account for the relevance of the cited documents to the purpose of the citing document.

Apart from evidential relations, another thirty or so structural relevance relationships are discerned by Green (1995b). Among them is also a metaphorical type of relation, in which "one complex unit is perceived as being structurally equivalent (as a whole or in part) to another" (op. cit., p. 660). This kind of relationship will recur in our case-study under the category of comparison.

12 The argument here is obviously not a deductively valid inference, since other reasons than that of the hypothesis could possibly be found for the domination of yellow and white flowers in the Northern Hemisphere. Rather, the argument should be seen as an inductive inference, in which the hypothesis, that UV-absorption ability functions as a selective mechanism for the flora of the Northern Hemisphere, appears as the conclusion with the evidence invoked by the citations as premisses.
Now, having shown the existence of other kinds of topical relevance than topic matching, viz. hierarchical and structural relationships\textsuperscript{13}, is there any use left for topic matching at all? The concept, admittedly, is by itself problematic and imprecise. Usually, topic matching is understood as meaning a correspondence or identity of descriptive terms of a document and terms used in a search query. Descriptive terms, in Dialog for example, can be controlled vocabulary descriptors, or identifiers, subject headings, classification codes, terms from document titles and/or abstracts etc. From this enumeration already it is evident that topic matching is a relative concept, that is dependent upon which database or information system and what query terms are used, fields searched, indexing depth etc. etc. This dependency may make it less successful as a search strategy in certain cases, particularly where the indexing terms of the system are quite specific, whereas the query is more general or cross-disciplinary.

As a possible illustration of such a case, consider the following example, originally provided by Garfield (1979), with the somewhat different purpose of showing the usefulness of citation indexing. The case concerns a patent granted to Schoeller in 1934 on "Wetting and Dispersing Agents for Use in the Textile Industries", classified by the Patent Office under 260-458, carbocyclic or acyclic-carbon compounds, which are acyclic sulfuric acid esters, in Chemical Abstracts (CA) indexed under "dispersing agents" and "wetting agents". In 1948 Lerner was granted a patent on "condensation products of cholesterol esters with polyethylene glycol & process for producing the same". That patent cited Schoeller's work, but was classified under 260-397.2, which covers sterols (including vitamin D), not corresponding to any of the classifications in which Schoeller's patent was placed; in CA it was indexed under "cholesterol esters" and "glycols, polyethylene", but not under "sterols". Also in 1948 a patent citing Schoeller was granted to Brown, on "waxy polyol ether-esters", classified under 260-234, carbohydrate esters. It was cross-referenced to 260-210 and 260-410.6, subclass 210 being glycosides, which is not applicable to either the Schoeller or the Lerner patent. Subclass 410.6 covers "synthetically produced higher fatty esters with acyclic polyoxy alcohols", applicable to the Schoeller patent, but not to the Lerner patent. The Brown patent, according to Garfield, was extensively indexed in CA under "waxes," "ethers," "esters," "polyol," "hydroxy compounds" ... and a number of organic compound headings. Now, Garfield states, the records of the Lerner patent show that the examiner did not find the Schoeller patent in the search of the patent file because he did not search any of the classes to which it was assigned. He found it in his files or his memory or elsewhere. But he found it, and wrote as a result, "Schoeller

\textsuperscript{13}One may of course ask, why insist that these relationships still be called topical relevance relationships at all? An intuitive answer would be that they are still principally related, as is topic matching, to the contents of documents. On the other hand, they also have functional aspects that appears to be closer to those of user centered concepts of relevance. For example, a piece of information from a source that, given a certain scientist's stock of knowledge at hand, is considered as having low credibility (or epistemic probability), is not likely to be invoked as evidence in an argument, no matter how well it would "fit" with the other premisses.
describes condensation of various acids, including stearic acid and its ester, with polyethylene glycols. "... These comments show that the relationship between the three patents, then is the condensation of acids with polyethylene glycols. (op. cit., p. 39)

The point seems to be that a search strategy based on topic matching would not have managed to discover the relationship between the three patents. But, what if a topic matching had been attempted using the term esters only, apparently present in the classifications of all three patents? Supposedly, such a search would simply have yielded too many non-relevant answers to be convenient (that is, the precision would be very low), since esters are used for so many totally different purposes, e.g. as aromatics, as wetting agents in textiles, as softening agents in PVC and rubber.14 In the case of patents it seems important to assure very good recall, in order to avoid legal suits for infringement on others rights, while retaining reasonable precision, to avoid information overload. Topic matching would seem to be ill-suited to meet the requirements of a good search strategy in cases like these.15

A further argument against topic matching as a method of finding relevant documents may be adduced if we recall again the "logical view" of relevant information as providing evidence for a hypothesis or conclusion, keeping in mind that two different pieces of information that jointly provide evidence to a case may often come from entirely different documents, the topics and indexing terms of which are also entirely different. As an example, consider again the extract from Utech (1975), quoted above. The fact that the citations serving as evidence in that passage come from entirely different subject areas makes it improbable, that the cited documents would have been retrieved by a single topic matching search.

Despite the seemingly abundant evidence against topic matching, there is a methodological problem here. Given the relativistic character of the concept, its dependency on, among other things, search query terms used and index terms assigned to documents, there is really no point in trying to prove the failure of topic matching to account for actual relevance relationships, since, in principle, it could always be proved to be insufficient in that respect, given sufficiently broad query terms and / or index

14 A search in the electronic version of Chemical Abstracts, CA Search database (Dialog file no. 399) reveals, admittedly, that in this case, since the patents are of rather old date, the fear that the search term esters would bring to much noise may not be wholly warranted. There are in fact only 7 records that have been assigned the descriptor "esters" for the time period 1930-1960. However, if the search is made to cover instead the time period 1930-1970, there are 15878 records that match, already a far to large number to browse through. To date (August 1, 1996) there are 206715 records matching the search query "esters/de" in the CA Search database.

15 Cf. also Brooks (1995), who showed that topical and broader descriptors are matched to records more easily than narrower descriptors, and Saracevic (1988) who showed that a vague query, with low subject specificity, but high complexity increases the probability of finding many items relevant to the query.
terms. In particular, when determining the relationship between cited and citing documents, any two documents could, in principle, be made to match. Therefore, in order to make sense of the claim that topic matching is sometimes ill suited to account for relevance relationships expressed by citations in scientific articles, the conditions under which a topic matching relationship is said to obtain must be specified in advance.

Thus, in this thesis henceforth by topic matching will be meant that two documents have least two descriptors in common, as indexed in a given scientific database. The database chosen for the case-study is Aerospace database (Dialog file no. 108). Note that in this restricted sense, while there is a clear limit for when and when not a topic matching relationship will obtain, levels of matching are allowed for. To express this relationship in quantitative terms, the so called union overlap measure, will be used, to be explained in section 4.2, dealing with the methodology used for the case-study. Hereby, it must be noted, the topic matching concept of this thesis differs substantially from that employed by Green (1995a,b), where it appears to stand for an all-or-nothing relationship: either a topic matching relation obtains or it doesn't. The purpose for which the concept is employed in this study admittedly may not be entirely compatible with that of Green (1995); here it is used to describe relationships between documents, viz. scientific journal articles, whereas in Green (1995b) it is meant to stand for a relationship between a user need, as expressed by a concordance to a religious document, and different topics of that same document. However, it seems to this author, that despite the valuable and largely successful effort to demonstrate the importance of topical relevance relationships other than topic matching, Green (1995a,b) actually fails to distinguish topic matching clearly from these other topical relevance relationships. For example, some adjectival and adverbial relations such as magnitude and temporal conditions are classified as topical matching relationships, while at the same time they are allowed to appear also as structural relationships (op. cit., p. 659).

An original and somewhat different approach to topical relevance and to relevance in general is found with Alfred Schutz, philosopher, phenomenologist and disciple of Edmund Husserl. Schutz wrote already in the 1940's a theoretical work on relevance, published posthumously in 1970, in which he distinguished three different types of relevance, that he considered to be simultaneously present in our consciousness: topical, interpretational and motivational relevance. We will not go into the details of Schutz's complex and highly interesting theory of relevance, but merely note some of its features that seem to have particular bearing on our case study.

To describe how these three aspects of relevance, topical, interpretational and motivational are interconnected, Schutz uses Carneades' classical example of someone confronting an unknown object in the dark corner of a room upon entering. Now, he asks,

Why, first of all, does this object appeal to his attention to such an extent that he makes it the theme of his interpreting activity? What makes the interpretation of this object at all a problem to him? There may be many other objects in this room, perhaps in other corners, perhaps in the same
corner of the room, which leave him entirely indifferent; yet they are all within his unstructuralized field of vision. They remain in the horizon. They do not become, in his actual situation, thematic. They present no problem, interpretative or otherwise, and he thus does not "pay attention" to them. This something which he will attempt to interpret either as a pile of rope or a serpent is, in some way, privileged among all the other objects in the room. It stands out over against them, it is from the outset - to introduce this term by anticipating later results - relevant to him.

(Schutz 1970, pp. 24f)

This relevance, according to Schutz (ibid.), is situational in the sense that it is dependent on the man's previous autobiography (including his stock of knowledge at hand, one might suppose), and the antecedent genesis of the actual situation (involving also, it might be suggested, the selection process leading to his entering the room at that very moment). These, in turn, govern his expectations upon entering the room. Assuming, for example, that the room is his own

He then expects upon entering to find his room more or less as he left it ... This set of expectations may constitute the unclarified but pertinent frame of all possible experiences of the the room he expects to have when entering at this specific moment. But this set of expectations, this field of the unproblematic which constitutes or at least coconstitutes the frame of reference of all possible experiences he expects to have, proves to be broken asunder by a novel experience having neither the mark of the unproblematic nor that of familiarity. ... the collapse of his expectations, the unexpected change, imposes upon him a change of his thematic field.

(ibid.)

It is interesting to note in passing how the last statement in the quoted passage seems to point in the same direction as Saracevic' concept of relevance, as something involving a change in the destination file of the user, here the mind of the man entering the room.

Topical relevance, then, is defined by Schutz (1970) as that by virtue of which something is constituted as problematic in the midst of the unstructuralized field of unproblematic familiarity.... But to make an object a problem, to make it the theme or topic of our thought, means nothing else than to conceive it as a dubious and questionable one, to segregate it from the background of unquestionable and unquestioned familiarity which is simply taken for granted.

(op. cit., pp. 26f)

Thus, if Schutz is correct, topical relevance would seem to be dependent on our expectations of the order of things. In a scientific context these expectations in turn would seem to depend on the hypothesis or theory that is guiding a scientist's research. The theory tells her or him what to expect as normal, if the theory is correct, and what will appear as problematic and, hence, in need of an explanation. But not only that, the theory, as part of the stock of knowledge at hand will also guide the scientist in the
interpretation of the, topically relevant, perceived phenomena; i.e. it will determine what is interpretatively relevant.

As an illustration of how expectations grounded in theory may influence the interpretation of scientific data consider the fact that the discovery of the "ozone hole" over the Antarctic, finally made public by Farman (1985), was actually delayed for some years, due to a mistrust in the extremely low values shown by ozone measurement data both from a team of British researchers using a traditional Dobson meter, and from the Nimbus 7 satellite. Since data did not fit predictions made from existing theoretical models and since they showed values considerably lower than ever before observed, it was assumed that they were the effect of bad measurements, and it seems that for some time they were regarded as more or less irrelevant (cf. Nolin 1995, pp. 119ff). In this case, then, expressing the situation in Schutz' terms, we would say that the measurement data, that were to stand stand out as "problematic", were initially judged to be of topical relevance to the purpose of describing the distribution of ozone in the Antarctic region.

If this had not been so, presumably no measurements would ever have been made. However, when these same measurement data, in view of the theory at hand, did not conform to what was expected, they were judged for some time to be interpretatively non-relevant to that same purpose. That is, rather than imposing an immediate change in "thematic field", as in the case with the man finding an unexpected object upon entering his room quoted from Schutz above, there was a conscious choice to ignore the data. In fact, since data were considered to be unreliable, new instruments were acquired by Farman and his team, in order to replace the old data with new, through what was believed to be better measurements (cf. Nolin 1995, pp. 119ff). Nevertheless, they were still retained in the background of the thematic field as one the possible alternatives to be considered. And when eventually theory changed, through the introduction of a new element - that of "polar stratospheric clouds", the unreliable, "faulty" data again became both topically and interpretatively relevant to the purpose developing the new theory (op. cit., pp. 122ff).

To describe, finally, how, topical and interpretational relevance are interconnected with motivational relevance, Schutz again invokes Carneades' example of a man confronting an unknown object in the dark corner of a room. Now, suppose the man happens to be afraid of snakes, maybe as a result of an early childhood experience. His fear of snakes is then "motivationally relevant for considering this object as dangerous, provided it is a snake, and for projecting its removal with a maximum of safety" (Schutz 1970, p. 50).

In order to find out whether the perceived object is in fact a snake or, e.g. a harmless pile of rope, the man must, however, make use of interpretatively relevant information, gained through further observation (e.g. watching the object starting to move) or by means of an experiment (e.g. hitting the object with a stick). Thus, Schutz continues, my fear of snakes is also mediately motivationally relevant for the limits up to which I have to follow the intrinsic interpretational relevances in order to be satisfied with the plausibility-degree of the outcome of this interpretation. But not only is this the case, for my fear of snakes might be motivationally relevant for the establishment of the either-or alternative,
that of "problematic possibility": "this is either a pile of rope or a snake," instead of: "this is either a pile of rope or a pile of clothes." Finally, my fear might even be motivationally relevant for the system of topical relevances for focusing this and no other object in the dark corner as worth my attention. (ibid.)

Summing up, then, both interpretational and topical relevances are dependent upon the motivational relevance as for (1) limits and scope of interpretation, (2) alternatives to be considered and (3) focus of attention - i.e. topic in a narrower sense.

All these three aspects of relevance will recur in the citation analysis on different levels. The purpose of this study is not to present a full-fledged interpretation and application of Schutz' theory of relevance on contextual citation analysis. Nevertheless, the citation functions described in the case-study would seem to involve elements of both topical, interpretational and motivational relevances in various proportions. An obvious example is the category motivation (see section 5.3 below), which is used in this thesis to describe citations serving as a motivation, in adjunction with other reasons, for the research reported by the citing authors or, more generally, for their writing their paper. (The motivations for writing a scientific paper may of course be manifold, and partly extra-scientific. The kind of motivations covered by the category motivation in this thesis, however, are essentially intra-scientific in character; i.e. they belong to the realm of scientific discourse and are stated openly, sometimes by means of citations, in scientific papers. Cf. also the distinction between citation functions and citation motivations introduced in section 3.2 below.). Another example might be the result category (see section 5.3 below), which is used here to describe a conditional relationship between cited information and citing source, typically occurring where the cited article belongs to another field of research than that of the citing source. In such instances, it would seem that the relevance of the cited information to the citing source can sometimes be described as that of serving as a key for the interpretation of data. Consider as an illustration of this again the case of Utech (1975), referred to earlier in this section, in which information about the distribution of ozone and UV-radiation served to suggest, in conjunction with other background information, a certain interpretation of the fact that yellow and white flowers are predominant in the Northern Hemisphere. Thus, the cited information appears to have been interpretatively relevant to the purpose of explaining this fact. Similar connections between Schutz' different aspects of relevance and the other citation functions described in this thesis could probably be found, for example, in cases where the method of a cited paper is criticised for yielding data that are insecure and unreliable, something that would seem to indicate a negative interpretational relevance.

The relationship between the topical and the interpretational relevances could also be expressed, in a very schematic and crude way, by means of the "infological equation",

\[ I = i (D, S, t) \]
where \( I \) is the information obtained through the interpretation \( i \), applied to the set of data \( D \), in view of the stock of knowledge \( S \), during a given time interval \( t \). (Langerfors 1973; adapted from Malmsjö 1995, p 39f).\(^{16}\) Thus, the interpretation acts upon a set of data and a stock of knowledge at hand, from which derives also the topical relevance of those same data. The time element appearing here is no doubt also important for the interpretation, as seen from the preceding example, not only as regards the length of the time interval during which the interpretational process runs, but also as for the specific point in time in the development of a field of research when the interpretation is made. Thus, the relevance judgments reflected through citations to scientific journal articles are likely to be affected not only by the theoretical approach of the citing author(s), but also by the time at which the citation occurs.

Before engaging in the citation analysis we must, however, touch briefly upon the question of why most of the information that surrounds us is actually considered not relevant to our purposes.

### 2.3 Non-relevance

Given the multitude of views of relevance it is clearly impossible to enumerate all the reasons that may be adduced for the non-relevance of a piece of information to a purpose. A few principal factors that appear to have a decisive influence on judgments of non-relevance may, however, be discerned. We have already seen one example of how theory and the consideration of the source of information (or data) as insecure led the scientists to disregard, for some time, the potential information contained in the data as irrelevant. (Cf. again the case of the Antarctic ozone hole referred to above; Nolin 1995, pp. 119ff). In such a case, though, it seems probable that a judgment of non-relevance would be tantamount to a decision not to cite the source of information considered as insecure. In fact, Farman, the leader of the British team of researchers who discovered the Antarctic ozone hole, deliberately delayed the publication of their measurement results and actually persuaded a young member of the team not to cite the results in his dissertation (ibid.).

Other factors influencing judgments of non-relevance are those connected with the stock-of-knowledge at hand, as emphasized by the "pertinence view" of relevance (cf. above section 2.1), e.g. actuality or recency of the information, accuracy, completeness etc.etc. However, it can be assumed that documents judged non-relevant on such grounds are in most cases not likely to be cited.

\(^{16}\) The "infological equation" is a convenient way of expressing the variables entering the transfer or creation of information; it does not, however, say much about the nature of the relationship between the elements involved, especially not about the relation between data, stock of knowledge and time. This objection is part of a wider discussion about the "theory-dependence" of data and already of observation, inherent in Schutz's theory.
What interests us here is, rather, those cases in which information is cited, in spite of its apparent non-relevance to the purpose at hand. One possible such case is that in which there appears to be redundancy among cited sources, i.e. where several sources containing essentially the same information is cited simultaneously. In any event, cases of this kind - and they are not few, will obviously fail to qualify for the logical concept of relevance employed by Cooper (1971), according to which relevant information must be part of a minimal premiss set, from which the answer to a search question can be deduced. However, there may in fact be several good reasons to cite apparently redundant information. For example, in order to add qualities of literary style or pedagogical force to a text. True, such motives are not likely to be prevalent in scientific discourse. But, there are other reasons as well why citations to apparently redundant information cannot so simply be dismissed as not relevant. Keeping in mind that most inferences in scientific research are in fact inductive, rather than deductive, the proliferation of citations to sources of essentially the same content can often be seen as a way of adducing evidence in support of a point that the author of a scientific article wants to put forward. This is a perfectly legitimate and rational way of reasoning; it is a well known principle in judicial matters that the joint testimony from several, independent witnesses has a stronger evidential force, than that produced by only one witness alone. In scientific articles this summoning of several witnesses in order to state a case sometimes takes the form of an enumeration of examples of previous research, intended to illustrate a point in the argument put forward. There are several such instances in the material of the case-study, classified under the category example (cf. below section 5.1).

Thus, multiple references to papers of essentially the same content should not always be seen simply as the result of a wish "to keep everybody happy' in the game of priority hunting" (Moravcsik 1975, p. 90). At least in some cases where cited information initially appears to be redundant, and, hence, not relevant, a closer analysis may reveal that the joint testimony of citations justifies them as having, nonetheless, some kind of evidential relevance to the purpose of the citing article. Other cases in which clearly non-relevant information is being cited, however, cannot be explained away that easily. Such cases are those where the motivations for citing are connectional, rather than professional, to use the terminology of a distinction introduced by Vinkler (1987, see: Liu 1993, p. 395). That is, they are citations motivated e.g. by personal relationships, prestigious reasons and the like. Luckily, these seem to account for only a small minority of all citations, according to the results of Vinkler (ibid.), which will be presented in brief in section 3.2, dealing with the various functions of citations. But first a general introduction to citation analysis is required.
3 Citation analysis

Citation analysis is a broad term that could be used to cover a variety of approaches to a given body of literature, whether scientific or not. First, it seems possible to distinguish both a quantitative and a qualitative aspect of citation analysis. Presumably, one could argue that a non-negligible part of the work performed by critics and theorists in the humanities might well be termed citation analysis, that is, an analysis undertaken in order to trace the influences exerted on a certain author by another, or several other authors. However, in fiction literature, at least, overt citations are rare, so maybe this type of research should better be termed text analysis. The kind of citation analysis undertaken in this study, although concerned with articles in scientific journals, bears a certain resemblance to the literary critics work, in the sense that it is qualitative rather than quantitative in character. What will be analyzed is a definite, limited number of texts, with the explicit aim of uncovering certain relationships between these texts and a selected pair of other texts. Therefore, in the historical review of citation analysis that follows the focus will clearly be on this latter kind of contextual or content-oriented citation studies. No claim is made to give an in any sense complete chronicle of previous research in the field; readers interested in a more comprehensive review are referred to the works of Garfield (1979) and Liu (1993).

3.1 Historical background

Citation analysis as a method of studying the scientific process developed mainly as a result of the invention of citation indexing. Citation indexing, in its turn, grew to a large extent out of the attempts made after World War II to construct automated indexing systems. Garfield (1979, pp.6ff) retells the interesting story in some detail. The oldest citation indexes, however, may in fact date way back in time. Some Bible concordances actually have properties resembling those of a citation index. That is, they permit a potential user to find, for any given Biblical verse, all the places in the rest of the Bible where this particular verse is 'quoted' - directly or indirectly.

Later in history, the usefulness of a citation index was discovered among professional lawyers. Shephard's Citations, founded in 1873, soon became an indispensable tool for searching legal decisions, providing citations to precedents used in the cases decided by federal and state courts and various federal administrative agencies in the USA. Initially, however, it was not an index to legal periodical literature as it is today. Nevertheless, it has served as a model for how a citation index could be organized in order to function as an effective search tool (op. cit. p. 7).

The Science Citation Index (SCI), the first annual edition of which was published in 1963 (op.cit., p. xiii), grew out of a series of three pilot citation indexes to the genetics literature. The project was originally financed by a grant from the National Institutes of Health, but it was carried through by the Institute for Scientific Information (ISI) in Philadelphia. The process of production began by the compilation of a multidisciplinary
database, which consisted of citations of all the material published in 1961 by the 600 journals covered by Current Contents and of all the references in that material. Then, in a second step, the citations that applied to genetics were extracted from this database to produce the 1961 Genetics Citation Index (op. cit., pp. 12f). According to Garfield (1979) "the pilot study that produced the genetics indexes showed a citation index to be eminently useful" in the sense that it was "deeper, more specific, and considerably broader in scope than conventional subject indexes. These qualities made it a more efficient and productive search tool." (op. cit., p. 15f). In particular, it seemed that the advantages of a citation index over traditional subject indexes "grew considerably more pronounced as the coverage of the index increased" (ibid.). For this reason, when the American government refused to finance further a publication of the multidisciplinary database, ISI decided to go along with the production on their own. The SCI was born. Already from the onset, the production was computerized - otherwise it would hardly have been feasible, considering the workload involved. This fact probably facilitated the transfer from the paperbound editions of SCI, and its younger siblings Social Science Citation Index (SSCI, from 1973) and Arts and Humanities Citation Index (AHCI, from 1978), to the online- and CD-ROM-versions of today.

Most of the work in which citation analysis has been applied to science, making use of SCI and SSCI, has been quantitative in character. Citation counts have been made, for example, in order to identify the core literature of a scientific discipline, to single out the most important journals in a field of research. Citation analysis has also been used as a method of research in the modern history of science; a notable example of the kind is a case study of the discovery of the structure of DNA (cf. Garfield 1979, pp. 83-9317). The method of co-citation clustering has been used for mapping the structure of

17 Garfield proves citation analysis to be a powerful tool for the unfolding of a historical development in science. However, it should probably best be seen as a complement to traditional historiographic methods and to text-analysis. In the case of the DNA-theory, for example, it seems that citation analysis failed to identify some of the important influences behind Watson's and Crick's famous 1953 paper, e.g., those of J.D. Bernal, I. Fankuchen and Gerhard Schramm. (Cf. Watson 1978, pp. 91ff). Also, the direct influence of Hershey and Chase, through a personal letter, on the work of Watson and Crick (op. cit., p.97), is not reflected in the network diagram of the development of the DNA-theory produced by the citation analysis undertaken by Garfield (op. cit., p. 88), where there is no link, neither direct nor indirect between these scientists. Most striking, perhaps, is the fact that only one direct citation to the famous Watson and Crick (1953) paper was identified as being of importance in this development.

Case-studies like these may be of value also for studying relevance relationships, although the objective is somewhat different from that of history of science. For example, Garfield dismisses some of the seemingly important citations to papers on methods and a reaction, as "probably ... not sufficiently critical to the development and proof of the DNA theory " (op. cit., p. 92f), since they were not mentioned in the account given by Isaac Asimov (The Genetic Code. - New York : New American Library, 1963). But, clearly, since these papers were cited by highly acknowledged scientists, who occupied themselves central positions in the development of the DNA-theory, they
scientific disciplines and for identification of specialties within disciplines (op. cit., pp. 99-141).

One of the pioneers in applying a qualitative approach to citation analysis was Lipetz (1965), who seems to have been the first to investigate the relationship between cited reference and citing document, with the explicit purpose of improving the selectivity of citation indexes. Lipetz’ system had four major groups, comprising twenty-nine different categories. The eleven categories of the first two groups, however, described the scientific contribution of the citing paper independently of the cited paper. Citation relationships, in a proper sense of the word, were thus expressed by the eighteen categories of group three and four, describing the "identity or continuity relationship of citing paper to cited paper" and the "disposition of the scientific contribution of the cited paper in the citing paper", respectively (op. cit. p. 83). Categories of group three were, for example "erratum" "continuation" and "inclusion", whereas group four comprised such categories as, e.g., "noted only", "applied", "affirmed" and "refuted". Obviously, the twenty-nine categories proposed were not intended to constitute a final judgment on the matter, since already at the outset Lipetz made the following very wise declaration:

Just as there are no absolute and exclusive terms by which a document may be fully described in subject indexing, there are no absolute and exclusive categories which fully describe the relationship of a citing publication to the cited publication. Relationship categories, like subject descriptions, may be multiplied, modified, qualified, or subdivided to infinity. (ibid.)

Since then, numerous attempts to formulate classification schemes for citation relationships have been made. Some twenty different such studies are reported by Liu (1993) in an extensive review, that will be our main reference and point of departure when we turn now to a discussion of the different functions of citations.

### 3.2 Citation functions

The multitude of schemes describing the functions or motivations of citation that have been proposed reflect various research objectives, as noted by Liu (1993):

1. to enhance the value of citation as search tools by testing the reliability of citation indexing

were in all likelihood judged relevant to this development at least by that point in time. For that very reason, they would at least appear to be worthy of a closer examination, aiming at a deeper understanding of this possible relevance relationship. In general, there sometimes seems to be a tendency to underestimate the importance of methodology for the development of science. In the long run, of course, only time will tell which theories, which inventions and which papers will prove to have been crucial for the advancement of science. The information scientist can learn from history, to some extent, but our concept of relevance must needs be based on the knowledge we possess today. Citation analysis may be one important tool in explicating this concept.
2. to describe the manifest roles and functions of the citations in certain subject literatures;
3. to assess quality of citations and clarify citation context
4. to define specific concepts attributed to the cited work by citing author
5. to look at the underlying motives for citing

(op.cit. p. 379f)

The last of these objectives is most prominent in citer motivation surveys, in which authors of scientific papers are asked questions about the motives for citing (and for not citing) this or that source. Liu herself is evidently a partisan of these kind of studies. They will, however, not concern us so much here. The reason for this will be discussed in section 4.2, where the methodology of the case study is described in detail. Thus, we acknowledge a fundamental distinction between, on the one hand, citation functions, describing the manifest use of cited information and, hence, the relationships between cited and citing documents, and on the other hand, citation motivations, describing rather the opinions of authors of scientific papers about the reasons behind the citations occurring in them. This study is primarily concerned with citation functions.

Now, it seems that research objectives 1-4 above may occur in various proportions in studies of both types, but perhaps more frequently in those dealing with citation functions. The predominant aim of this study and of most of the studies referred to in what follows is obviously the second on the list, "to describe the manifest roles and functions of the citations in certain subject literatures".

On a more general level, the various normative or sociological functions of citations have been summarized by Mitra:

They contribute to scientific bricklaying, provide the necessary 'currency' for recognition of a particular scientist among his peers as well as for establishing his property rights and priority claims with respect to the scientific contribution he makes, constitute an important source of information for the scientist, serve as a tool for some... techniques of literature indexing and... help to gauge the information gathering habits of scientists and indicate therefrom the literature that is indispensable to their work.

(Mitra 1970; quoted from Liu 1993, p. 372f)

Though the last statement may obviously be disputed - far from all citations appear to be "indispensable", the just quoted passage clearly indicates the overall importance of citation in scientific communication. However, the various classification schemes for citations that have been developed reflect this importance in different ways, "with each classifier regarding his or her problem as unique" (Liu 1993, p. 400). There may well be methodological justification for this, as Peritz (1983) seems to suggest, when summarizing the problems of context / content analysis of citations:

... a practical method of 'labelling' citation for retrieval purposes has not yet been found; the assessment of quality and context, let alone underlying motives, involves a large degree of personal judgement as well as an-in
depth knowledge of the subject matter; manifest roles of citations differ from one field to another and therefore universal classification systems seem inappropriate.

(op. cit., p. 303; quoted from Liu 1995, p. 402)

Among the classification schemes suggested there appears to be a division first between those that describe the strength of the citation connection between documents, and those that try to describe, rather, the nature of this relationship. The systems proposed by Moravcsik (1975), Chubin (1975) and Bonzi (1982), for example, seem to be of the former kind. Moravcsik (1975) introduced a model comprising eight categories arranged into four dichotomous pairs, viz. 1. conceptual vs. operational (distinguishing citations "made in connection with a concept or theory that is used in the referring paper" from those connected, rather, with "a tool or physical technique"), 2. organic vs. perfunctory (distinguishing citations that are "truly needed for the understanding of the referring paper" from those that are not necessary in the same sense), 3. evolutionary vs. juxtapositional (distinguishing cases in which the citing paper is "built on the foundations provided by the reference" from those that are, rather, an alternative to it), and, finally, 4. confirmative vs. negational citations, the latter term covering those cases in which the correctness of the cited paper is disputed (op.cit., p. 88). Apparently, the last three of these distinctions at least describe citation relations of different strength, with the first term of each pair expressing the stronger link.

The model of Moravcsik (1975) was modified by Chubin (1975), who retained the fourth of the above pairs as essential, while introducing two further levels in the second. In this way the connection strength of the affirmative (replacing the term "confirmative") citations was expressed by four different categories, from strongest to weakest link, viz. 1. basic (replacing "organic"), 2. subsidiary, 3. additional information and 4. perfunctory (op. cit. pp. 426ff). The negational references, on the other hand, were subdivided into 5. partial and 6. total, the latter representing those citations where "the citer refers to the paper as being completely wrong and offers an independent interpretation or solution", whereas the former represents those cases in which "a citer suggests that the paper is erroneous in part and then offers a correction" (ibid.).

The four category model used by Bonzi (1982) is somewhat different. It is expressly designed "to measure citation relevance" and is more pronounced in its quantitative approach. The four categories of this scheme are:

(1) not specifically mentioned in the text (e.g., "Several studies have dealt with..."),
(2) barely mentioned in text (e.g. "Smith has studied the impact of..."),
(3) one quotation or discussion of one point in text (e.g. "Smith found that...")
(4) two or more quotations or points discussed in text.

(op.cit., p. 211)

Although important, the division between "quantitative" citation classification systems, like those just presented, and those, on the other hand, more aimed at describing the
qualitative nature of citations, should not be considered as irreconcilable. On the face of
it, at least, the four "measures" introduced by Bonzi (1982) above, could easily be
combined with a qualitative scheme, using categories such as "evidential", "methodological" and the like. In particular, the classification scheme employed in the
case-study performed for this thesis, although essentially qualitative in character, has a
separate category for quotations, and the number of references to a cited text, in case
there is more than one, is also taken account of. As for the systems of Moravcsik (1975)
and Chubin (1975) it would seem possible to "translate" at least some of their categories
into a more qualitative classifications. For example, citations judged to be only
"perfunctory" could in many instances perhaps be seen as providing background
information, whereas those regarded as "organic" or "basic" would be more likely to
serve as evidence. In like fashion, "partial negational" citations could perhaps just as
well be described as instances of modification, whereas "total negational" citations, if
they ever occur (no instances were actually found by Chubin 1975), could possibly be
expressed by a category of objection. (Words appearing in block letters in this
paragraph and the following are categories used in the case-study of chapter 5).

Among "qualitative" citation classification systems there is a great diversity, which
sometimes makes it difficult to compare results of different studies. This is not to say,
however, that there are no similarities between some of the proposed systems. Several
authors, for example, include among the categories describing citation functions those
of providing background information (Garfield 1979, Duncan 1981, Peritz 1983 etc.)
or identifying methodology (Hodges 1978, Finney 1979, Garfield 1979, Duncan 1981,
Peritz 1983 etc.). Although the exact labels used for classifying citation relationships
may differ, it also seems clear that they often cover cases of essentially similar type, as
for example the following labels used for describing some kind of evidential citation
relationship: prove, verify, substantiate (Spiegel-Rösing 1977), evidential (Hodges
1978), confirmation (Finney 1979), support evaluative opinion / factual statement
(Frost 1979), substantiating claims (Garfield 1979), corroboration (Duncan 1981) etc.

There would be no point in reproducing here the different classification schemes that
have been proposed over the years, many of which are listed in Liu (1993). Suffice it to
say that most of the categories used for classification of citations in the present study
(see section 5.1) have been found to be represented in the literature of citation analysis.
This is perhaps not surprising, given the volume of this literature. It should be noted,
however, that the categories employed in the present case-study were developed
essentially from the material at hand, that is, from actual citations in the articles studied,
before getting acquainted with other citation classification schemes. Thus, the obvious
resemblances between the present scheme and other such systems may be seen as an
indication that the number of functions served by citations in scientific articles is rather
limited, ranging somewhere between, say, ten and thirty. The appropriate number must
be determined from case to case and with respect to the purpose to be served by the
classification system. If the aim is primarily that described by research objective 2
above, a larger number of categories may be permissible, in order to describe properly a
certain body of literature. But, if the purpose of the classification scheme is rather that of
enhancing citation indexing as a tool for information retrieval (research objective 1
above), demands for uniformity and ease of use will be higher, and consequently the plausible number of categories will probably be lower.

Among the studies concerned with citation motivations, rather than with functions, Vinkler (1987) stands out as one of the few that investigated also the reasons for omitting citations, which will be dealt with in the next section. Introducing a distinction between professional and connectional motivations for citing, the former being linked to the theoretical and practical aspects of an author's research, whereas the latter are related to personal, social or other external factors connecting citing and cited author, Vinkler found in particular, that only 2% of a total of 484 citations were made purely for connectional reasons, whereas 81% had predominantly professional motivations and 17% resulted from a combination of professional and connectional motivations (see: Liu 1993, p. 395). Vinkler further introduced the concept of citation threshold to describe quantitatively the strength of the motivation to cite, and his study confirmed the assumption that "the citation threshold depends primarily on the professional relevance of the work potentially citable in the given paper" (Vinkler 1987, p. 69; quoted from Liu, ibid.). This result seems reassuring, in view of the attempt in this study at uncovering relevance relationships by means of citation analysis.

Before looking into the relationship between the function of citations and the concept of relevance in chapter 4, however, something must also be said about the reasons why most documents fail to be cited. Although no complete or generally valid answer can be given, this question will occupy us in the following section.

3.3 Non-citedness

There are probably at least as many reasons for not citing, as there are documents. Clearly, while working on a scientific paper, the overwhelming part of all the documents that are "out there" never even appear as candidates for citation. They are simply irrelevant for the purpose at hand. This section, however, is not concerned with them; it deals, rather, with some of the reasons for not citing another document, that does appear to be topically relevant to a scientific problem or hypothesis at hand.

First, we can distinguish a group of reasons that all have to do with limited accessibility of information. The most obvious case is perhaps where the document judged relevant is unknown by the potential user. (In this case, clearly, the relevance judgment must be made by someone else.) We may call this the case of non-retrieval or non-exposure; it means simply that the potential user has never been confronted neither with a bibliographical record of the document, nor with the document itself. The possibility of a non-exposure must always be taken into account in a citation analysis. For example, it may be that citation rates for two topically closely related and qualitatively equal
documents prove to be skewed, because of their appearing in journals with radically different circulation figures and / or impact factor.\(^{18}\)

Secondly, there is the case of a failure to acquire a document judged potentially relevant, due to obstacles of practical nature, for example. This, it might be suspected, is something that is more likely to affect students or non-professionals, and should be less common with qualified scientists, working at institutions with well equipped libraries. However, research has shown that scientists, in sharp contrast to humanists, depend mostly on their personal collections for their information provision; 73.9% of citations in published works by scientists are to be located in their personal collections, compared to 35.8% of citations by humanists. (Soper 1976, see: Budd 1989, p. 9). That is to say, although relevant documents may in fact be both known and in principle accessible, in several cases they still will not be acquired, read and cited, depending on the personal habits of the user, or simply due to human inertia.

Thirdly, there is the possibility of an erroneous rejection of a document as not relevant, while on a closer scrutiny it would in fact have proved to be relevant. We may call this the case of failure of recognition. It is difficult to estimate how often this phenomenon actually occurs. It seems reasonable, however, to assume that it occurs more often with non-specialists, than with specialists in the field of research to which the problem to be solved or hypothesis for which confirmation is sought belongs. A specialist in the field should normally be in a better position to judge about the value of new information in the field and see if and where it "fits in" with other information.

The three reasons for non-citedness mentioned so far can be compensated for to some extent in a citation analysis, although there seems to be no way of escaping completely their influence on citation behavior. For example, to minimize the influence of "non-exposure" the documents selected for the citation analysis can be chosen from journals of equal or nearly equal prestige, accessibility, circulation and / or impact factor. Ideally, in a qualitative citation analysis like the one performed in this thesis, articles could be selected from one and the same thematic issue of a journal, in which case there would be a very high probability that someone citing one of the articles also had been exposed to other articles in the same issue of the journal. The second reason for non-citedness mentioned above could also be handled in this way. The third reason, failure of recognition, is perhaps more difficult to detect and to deal with. But at least, as suggested by what has already been said, it is possible to restrict the citation analysis to experts in a field of research in order to minimize the possible influence of this factor.

A further reason for not citing that appears in the literature concerned with citation analysis is the phenomenon of obliteration, occurring "when a scientist's work becomes

\(^{18}\)The impact factor was introduced into SCI to compensate for the inequalities in citation counts that result from unequal numbers of items being published by large and small journals. It is a measure showing "the average citation rate per published item. This is done by dividing the number of times the journal has been cited by the number of items it has published." (Garfield 1979, p. 24)
so generic to the field, so integrated into its body of knowledge that people frequently
neglect to cite it explicitly" (Garfield 1979, p. 247). The risk - or the advantage of being
obliterated is obviously unequally distributed among documents, as MacRoberts (1987)
observed, and there may also be considerable differences between different disciplines
in this respect. It is a well known fact that within the humanities documents tend to be
more long-lived than in the sciences. The potential for obliteration, thus, has to be
assessed individually in any single case. MacRoberts (op.cit.), examining articles in
the history of genetics, found that only thirty percent of actual influences were being
cited. The phenomenon of obliteration does not, however, have to bother us too much
here, since it is of importance primarily in quantitative citation studies, using citation
counts in an attempt to "measure quality or impact of work, to establish cognitive
pedigrees, or to map scientific networks or specialities" (ibid.). In this study we are
concerned not so much with overall citation rates, but rather with the nature of the
citation relationships of the documents examined.

We turn now to those cases where there is a conscient decision not to make use of
information that appears to be relevant. Wilson (1995), in particular, points to several
possible reasons behind a decision not to make use of relevant information, the focus
being more generally on non-use, though, rather than on non-citation. Among the
explanations suggested by Wilson are "specialization", "territoriality", the "safely
ignorable" and the "unmanageable" - all four meaning that information about "some
distinguishable field of phenomena" X, that is thought to be relevant to a a field of
phenomena A, which is depending partly on X, is ignored (1) in order to concentrate the
research efforts on other aspects of A ("specialization"), or, (2) because X is regarded as
someone else's business, i.e. as belonging to another specialty ("territoriality"), or, (3)
because X is considered unimportant, as having a small impact on A ("safely
ignorable"), or, finally, (4) X is ignored "because it is intractible, does not yield to these
methods, is beyond one's competence, understanding, resources, etc." ("unamanageable") (op.cit., p. 47)

Although these reasons for non-use of information may not be immediately
translateable into reasons for not citing, they indicate on a more general level some
possible explanations behind citation patterns, should the assumption of the theory-
dependance of relevance judgement, as expressed by citations, be corroborated. In

19 An indication can perhaps be given by the level of citedness of classics within
different fields. A search in Social Scisearch (Dialog file 7) revealed, for example that
Emile Durkheim's classical work Suicide was cited no less than 106 times during the
year 1995, possibly indicating a low impact of obliteration in this case, although the real
number of documents on which Durkheim's work exerted its influence - whether it was
cited or not, is of course difficult to estimate.

20 Should it be that the figures representing the distribution of theoretical approach of the
citing articles, that are presented in the case-study, proved to be skewed when compared
with the actual number of "influences", whether cited or not, this would in itself be an
interesting fact, possibly tending to strengthen further the hypothesis of citations as
being "theory-dependent".
particular, a combination of these reasons for ignoring information would seem to offer a probable explanation of the fact that one of the theoretical approaches described in our case-study can be seen to ignore a certain type of data, possibly in order not to make mathematical and chemical formulas unduly complicated (cf. Nolin 1995, p. 60). We will return to this point in chapter 5, in the introduction to stratospheric ozone monitoring, and in chapter 6, when discussing the possible relevance of **background** information.

*Vinkler* (1987), as already mentioned, deals more specifically with reasons for not citing, rather than the more general non-use. The primary reason for omitting citation was found by Vinkler to be, perhaps not surprising, that information was "professionally not relevant" (cf. Liu 1993, p. 395). Other motivations for not citing included: taking common knowledge as one's own (this may be one possible cause of the phenomenon of *obliteration* described above), using review materials instead of original articles, artificial restriction of number of references (self-imposed or as decided by publishing journal), and "potentially citable work insignificant or not fully understandable" (ibid.).

The various possible reasons given above for not citing potentially relevant information should all be kept in mind as we now turn to investigate further the relationship between relevance and citations.
4 Relevance and citation analysis

4.1 The relevance of citations

Any research question regarding a phenomenon X, such as "What is the cause of X?" or "What are the properties of X?", presupposes already some previous understanding of X. This understanding may stem from two different types of information sources: from own observation, or from some kind of information record (electronic, print or other). In the latter case it seems natural to ask further, from where does this source have it's information. And so I am almost automatically led to check the references of my first source, the underlying assumption obviously being that these cited sources might contain some more information pertaining to my initial question, that is not recorded in my first source. This is a fairly simple, almost self-evident way in which citations may be used as tools for retrieving information considered relevant to my initial question. It is in fact a well-known, and widely-used search-strategy, to start out from a key-paper in a field of research, and continue by following up on those references of that paper that appear to be most pertinent to the problem at hand, then in turn following up on their references. This process is also known as chain-searching.

The idea behind citation indexing as a means to uncover relevance relationships, however, is to project this assumed relevance relationship forward in time from my original source, and examine later works, that cite my original source, in the hope of finding new information pertaining to my original question. But, whereas in chain-searching the decision whether to follow up or not on a reference is guided by the context in which the reference is made, in traditional citation indexes no such context is present. True, when conducting a search, e.g. in Scisearch, the electronic version of Science Citation Index, in most instances there is extensive information such as abstracts, descriptors, research fronts etc. given for each citing document to provide such guidance. All this information, however, may still not be enough to give the searcher a clear idea about why the original source was cited by the document described in the record under consideration.

Another observation that seems to speak against citation indexing as a means of discovering new, relevant information is that it is, to its very nature, rather conservative; it can only retrieve information that has actually been used already in a certain context. But, it's capability of detecting information of potential use, of discovering latent, hitherto unnoticed relevance relationships, and thus spurring novelty, appears to be small. A critique similar to this has been mounted by others, who claim that citation analysis is incapable of detecting premature discoveries that are still largely unknown to the scientific community. Garfield (1979, p. 247), one of the pioneers of citation indexing, has termed this kind of criticism the Mendel syndrome, since these critics often refer to the work of Gregor Mendel, which was ignored for decades by geneticists, although clearly relevant to the problems of their field of research. A possible response to this type of criticism would be simply to submit, arguing that, after all, "relevance means being part of the paradigm, or public knowledge, or consensus in a field"
(Foskett 1972, p. 78), and that citations can of course only record actual, not potential use of information. On the other hand, it seems questionable whether other traditional retrieval techniques, such as topic matching, would do much better in cases like these. Consider, for example, the case of Riemann's non-euclidean, elliptical geometry, which was long considered to have no possible applications, but which more than fifty years later proved essential in the development of the general theory of relativity. It is doubtful whether, in fact, any information retrieval system other than the mind of Albert Einstein would have been able to detect the relevance of Riemann's geometry for that purpose. Another example is provided by Swanson (1990), that of the discovery of dietary fish oil as a treatment for Raynaud's syndrome, a disease causing abnormally high blood viscosity and red blood cell rigidity in some patients. A search in MEDLINE and other databases revealed that the intersection set representing the topic matching of fish oil terms with those representing Raynaud's syndrome was actually void (op. cit., p. 32). A further search in SciSearch (Dialog file no. 434) confirmed the "bibliographic isolation" of the two bodies of literature: "With rare exception, they do not have authors in common; they do not cite each other; they do not cite the same literature", as Green (1995a, p. 650) comments. However, the two literatures are "evidentially relevant to the user's information need. But they are not on the same topic as the user need, that is, treatment options for Raynaud's syndrome" (ibid.). The lesson to be learned from this example, thus, is that where citation indexing fails to detect a relevance relationship, topic matching may prove to be no better as search strategy.21

Leaving the question of retrieval efficiency behind, let us consider instead, on a more theoretical level, the possible relatedness between citation and relevance relationships.

Figure 2: Relationship between relevance and citations

It would perhaps be tempting to describe the relationship between relevance (R) and citations (C) as in Figure 2, that is as a relationship between documents in an imaginary

---

21 On the other hand, one could possibly argue that both citation indexing and topic matching spurred the discovery reported in this case, at least indirectly, through the "negative searches" that were conducted in order to test the hypothesis. Cf. Swanson (1990, p. 32).
collection D. The shadowed area, \( R \cap C \), would then represent documents that are cited because of their being relevant to the apparent purpose of the citing article, while the areas \( a \), \( b \) and \( n \) would represent respectively documents that are relevant but are not being cited, documents that are cited while not being relevant and documents that are neither cited nor relevant. The reasons for the non-citedness of the documents in area \( a \) could be any of those enumerated in section 3.3, e.g. obliteration, not retrieved/unknown, plagiarism, political reasons, prestige, redundancy etc. In the same fashion, a number of possible motivations for the citations to the non-relevant documents in area \( b \) can be given, e.g. paying homage, borrowing prestige, political reasons, etc. The documents of area \( n \), those that are neither relevant nor cited, seem less problematic and need not concern us here.

The focus of our interest, thus, is the intersection - \( R \cap C \) - of relevance and citations. It is here that we expect to find evidence for the influence of "theory", in the sense of "metaphysical paradigm" referred to in chapter 1, on the selection of information for citation. At the same time, however, we have to allow for the possibility that some of the reasons for citing non-relevant documents (area \( b \) of Fig. 2), may in some cases co-occur with theoretical or other intrascientific reasons for citing. Thus, for example, it is quite conceivable that someone writing an essay for a Festschrift deliberately chooses to cite works of the person in whose honor the Festschrift is published as a way of paying homage, while at the same time, since the work of that same person may well be part of the subject of the Festschrift, these citations can be seen as perfectly relevant to the purpose of the essay. The example shows, that extrascientific reasons for citing do not always have to be seen as having an adverse effect on relevance and rationality in science. To be sure, numerous other examples of a more dubious kind could be adduced, in which it would not always be so easy to distinguish e.g. economical or political considerations from relevance judgments as grounds for citations.

The problem with the description in Fig. 2, however, is that one and the same document can be cited several times, even by the same citing article, and for many different reasons. At one time the motivation for citing may be based exclusively on relevance judgments, while at other times it may be grounded e.g. on political considerations that have very little to do with relevance. In general, relevance is neither a property of documents, nor a relationship holding between documents. Rather, relevance is better described as a relationship holding between the contents of documents and a purpose of use (or "information need"). And, given the fact that documents sometimes consist of many different, only loosely related parts (a model example would be the Bible), obviously one part of a document can be highly relevant to a purpose (or "information need"), while other parts of the very same document are not relevant at all.

Much the same goes for citations. As Martyn (1975) has pointed out, a citation is not a unit of information, but an "event". Although no specification is given as to what is to be understood by an "event" in this context, a possible interpretation is that citations must be viewed as something that is dependent upon the time and place of their occurrence, just as an event is something that always takes place at a definite point in
time and in a certain setting.22 The statement that a citation is not a unit of information, furthermore, points to the fact that what is being cited is in reality often not a whole document, but only a small part or a particular point of a document (cf. also Cano 1989, p. 289). The cited passage may be more or less closely, sometimes only very loosely related to the overall topic of the document being cited. Also, even in those cases where the citation is to a key passage of a document, one that appears to be closely related to its main theme, the citation of that passage may sometimes serve a completely different purpose, than that intended by the original document, as for example when the formula E=mc² is being used in commercials. Every instance in which a piece of recorded information is used by someone else than the originator may not be that extreme, but the likelihood that the purpose served by the information used will be exactly the same as in the original document appears to be small. In fact, even the exact copy of a whole document may not serve the same purpose as the original. People dealing in counterfeit art generally have something else in mind than the aesthetic impression gained from these works.

Thus, neither relevance relationships nor citations can be seen as consistent, discrete units. Consequently, they do not seem to behave as elements in a set, at least not according to the definitions of classical set theory, where one of the requirements on a set is, precisely, that every element should appear only once in the set. (Kamke 1950, p. 1).

Given the very complex nature of the relationship between relevance and citation, is there still any hope that a qualitative citation analysis could reveal anything that might be of interest for a reconstruction of the concept of relevance? If we are to believe Martyn (1975), there is:

At the very least, it is safe to say that when an item is cited, the citing author has heard of the reference and believes it to be sufficiently relevant to the topic of his work, or to a statement or remark contained in it, to be worth citing. The degree of relevance to the main theme of the citing paper may sometimes be inferred by examination of the citing paper but is not necessarily obvious from the bibliographic record of the occurrence of the citation. (op.cit., p. 291)

Another reason to retain confidence in the idea of citations as a kind of recorded relevance judgments appears if we recall for a moment Saracevic' explication of relevance as the measure of change effected when information is transmitted from one file to another (Saracevic 1976, p. 93). Nothing hinders us from regarding a text, a document as a kind of file. It contains information, it can be searched, it consists of separate segments that could be seen as records etc. A citation, then, is the result of a decision to include a new "record" in the file, thereby introducing a change. But again

22 A probabilistic interpretation of 'event' seems less plausible, since one of the requirements of an event in that sense is precisely that it can be seen as an element of a set, that is as a kind of unit.
we have to remember that such changes in files, manifested through citation behavior, may also have other causes than relevance considerations.

Finally, we may recall again the evidence produced by Vinkler (1987, see: Liu 1993, p. 395), according to which clearly relevance-based professional motivations accounted for 81% of the citations among 484 references analysed, while only 2% of the citations made for connectional reasons, the remaining 17% resulting from a combination of the two.

On the other hand, challenging this evidence (although not contradicting it outright), Liu (1995) reports from her earlier studies of the citing practices of Chinese physicists that "only a minority of scientists said that more than 80% of their citations were essential" (op.cit., p. 396). Summing up her results Liu concludes that "an author's citing behavior is unique, personal and complex" (ibid.). Liu's studies, just like Vinkler (1987), are expressively concerned with citation motivations, rather than with functions, using survey method's with direct questioning of author's of the citing papers. This poses an important methodological problem: Why should we rely on the relevance judgment of a citing author, more than on that of the reader? This question, although not answered conclusively, will be discussed in the next section, in which the method used for the case-study is presented.

4.2 Methodology

Liu (1993) in her review distinguished three basic methodological approaches to qualitative citation analysis:

1. classification of citation context;
2. citation content analysis;
3. citer motivation survey.

(op.cit., p. 379)

The methodology applied in this thesis can be said to be located somewhere in between 1 and 2, the aim being not so much that of testing the reliability of citation indexes, but rather "to describe the manifest roles and functions of the citations" (ibid.), although no claim is made that the results should be considered to be valid for the whole subject literature of the research field studied.

Relevance, as we have seen, has been expressed as a relationship holding between, on the one hand, a piece of information or an information source, and on the other hand, a user need. Now, the crucial question for the methodology applied in this thesis, is to what extent citations represent accurately this assumed user need. However, as Cooper (1971, p. 22) has observed, such an "information need" is itself only a theoretical construct, which serves to define the concept of relevance. Every representation of such a need must evidently be expressed as some kind of linguistic entity, just as the information assumed to be relevant to that need is also a linguistic entity of some kind. The present study is concerned with a special case of this relationship between
linguistic entities, viz. that between two texts, the cited and the citing article. The underlying assumption, thus, is that citation functions, derived by analyzing the content and context in which citations appear, can accurately represent this "information need".

As noted above, surveys or interviews with authors of scientific papers is another method that has been used for expressing this need. Such studies may have an interest as one way of describing how scientists actually work. However, the question must be raised: whose information need is it that counts? Or, to pose the question somewhat differently: who is the user? Is it the author of a citing text, or is it rather the reader, who decides whether or not to follow up on a reference, on the basis of the manifest content and context of citations in the text. It seems to the present author, that an analysis of citation functions is better suited than a survey method to represent this latter user need.23

Another reason for not employing a survey method for this study may be adduced. It has to do with the time aspect involved. As noted above (cf. also section 3.2), surveys have been used primarily to study citer motivations. Recalling for a moment the concept of motivational relevance, we may agree with Schutz (1970) that living in our action we have different interest in its motivationally relevant moments than when we merely project such an action, or when we look back at an action already performed in respect of its outcome or the single steps by which it was actualized and the project which preceeded it. In each of these instances, other chains of elements will prove to be motivationally relevant.

(op. cit., p. 52)

Now, obviously, this may pose a methodological problem. For example, the motivation for the initial selection of an information source for scrutiny may be entirely different from the motivation for citing it. Hence, a study based on interviews or a survey should have to specify at what point in time during the process of selection and writing the motivations given by the citing author were valid. And what guarantee is there that the motivations that authors give in their answers will not be rationalizations post factum, whether conscient or not, rather than true representations of the actual motivations at the time of writing?24

23 The survey method used by Cano (1989), although results are interesting, should be used rather to ask readers of scientific papers about the "utility level" of citations (perhaps comparing judgments with those of citing authors). There is no reason why the citing author should be in a privileged position to judge about the utility or relevance of citations; especially not if the study is made for the purpose of enhancing citation indexing as a tool for information retrieval.

24 As Chubin (1975) remarks, discussing the utility of data obtained by direct questioning of citing authors, "both the candour and recall of authors may be lacking, however, rendering such data impressionistic, selective and self-serving". This, of course, is not meant to imply that authors of scientific papers would be consciously lying about their motivations for citing. Rather, it is simply a recognition of the fact that
To be sure, there is a time aspect involved also in the analysis of citation functions. However, the impact of the time factor is not in the same degree as with surveys dependent on human memory and seems easier to control with this method, since to a large extent it is inherent already in the citing text, for example, in those frequent cases where a citation provides background information in an introductory overview of previous research in the field. In another sense, though, the analysis of citation functions depends only on one specific point in time for its validity. The citation function described, that is the claim of this thesis, is the function served by the citation in its context, i.e. at the date of publishing of the citing article. The analysis requires, it is true, some background knowledge of the state of the art of the cited and citing articles' field of research at this point in time.

The necessary background information for the case-study performed here, it was noted already in the introduction of the thesis, was provided by Nolin (1995), whose extensive account of the development in the field of research studied helped evaluate both the theoretical approach of cited and citing articles and the possible functions served by the citations. In fact, this work was instrumental already in the choice of stratospheric ozone monitoring as the particular field of research, from which the documents of the case-study was later selected. That is, it was realized already before the actual research began, that a contextual citation analysis of the kind performed here, would require at least some basic knowledge of the subject of the documents analyzed. Other studies of a similar kind, e.g. Harter (1993), have used documents from library- and information science for the analysis. This may appear to be a natural choice, since it is the field of research, in which information scientists have not only basic, but expert knowledge. However, it seemed more challenging to this author to make an attempt, for once, at breaking away from the narrow confines of one's own discipline, and venture out into the natural sciences. Libraries are supposed to serve all disciplines and all fields of research. But, if library- and information science restricts its object of study only to documents from within its own discipline, there is an obvious risk that the results of the research undertaken will ultimately concern only library- and information scientists, not the library users. Nolin (1995), employing an external, historical and meta-scientific perspective, made accessible to a lay reader a field of research within the natural sciences, situated at the crossroads of several disciplines, viz. chemistry, geophysics, meteorology and others, which made it seem an almost ideal object of study for a contextual citation analysis aiming at detecting a possible impact of differences in theoretical approach on citation behavior.

The specific methodology applied in the case-study of this thesis is further complicated by the investigation and comparison of citation functions with topic matching relationships. Green (1995b) observed the following methodological problem:

Traditionally, a study examining the topical relationships between user needs and relevant texts would make use of live reference data, comparing human memory is sometimes short, especially when it comes to our own previous inner states of mind.
the topics of statements of user needs with the topics of retrieved documents deemed relevant to those needs. The use of such data would tend to be problematic because the search result could be compromised by the very assumption of topic matching that the study is investigating: To the extent that searchers have been trained to look for documents whose topics match the topic of the user need, the predominant relationship between user need and relevant texts retrieved in actual searches will be topic matching. ... we suggest that relevant documents whose topics do not exactly match the topic of the user need will have disproportionate representation in the non-retrieved set. Thus, traditional data sources are likely to be biased and results based on them skewed.

(op.cit., p. 654f)

The "user needs" in this thesis, as noted, are those expressed by actual citations in context, and the "relevant texts" are the cited documents, the pair of cited articles selected for the case-study, which are assumed to be relevant to the function / purpose served by the citation, if no evidence is found indicating non-relevance. Such evidence would be constituted by indications of the kind that were discussed in sections 2.3 and 3.2 above, for example, the presence of personal ties between citing and cited author, while no other function of the citation can be detected than that of providing "redundant" background information, that is, information which is also provided by other citations at the same location. Thus, instead of following the strategy of Green (1995b), trying to bypass the problem stated in the quotation above by using a source of information that is not "biased" in this sense, the present study will regard the assertion that "the predominant relationship between user need and relevant texts retrieved ...will be topic matching" (ibid.) as an hypothesis which is put to test in the case-study...That is, the case-study by itself is expected to give an indication about the extent to which topic matching is indeed correlated with citation.

The method used for the case-study in the next chapter was presented briefly in section 1.1. In the following the steps involved will get a somewhat more detailed description, while the presentation of the material will have to wait until the next chapter.

1. The first step was thus the selection of two scientific articles from the research field concerned with monitoring of the stratospheric ozone layer. In order to minimize the possible impact of factors such as time, journal prestige, differences in exposure etc on citation patterns, it was required that both articles be published the same year, in equally reputed and generally read scientific journals. It was further required that a topic matching relationship should hold between the two articles; topic matching being defined for the purpose of the case-study as the property of having at least two descriptors in common, as indexed in one of Knight-Ridders DIALOG databases. A third requirement was that the two articles represent different theoretical approaches in the field of research studied. The choice fell on two articles published in 1974, in the multidisciplinary,well-reputed scientific journals Science and Nature. (The articles and the specific reasons for selecting them will be presented in the next chapter, sections 5.1 and 5.2)
2. The second step in the method applied was the retrieval of documents that cite the two selected articles. This was effected by means of DIALOG SciSearch, an online version of the Science Citation Index. For one of the articles, records of all the 55 citing documents were retrieved, saved and printed out. For the other, which has a considerably higher rate of citedness, a further selection was made by limiting the search to certain publishing years. The selection of publishing years was done both with respect to the historic development within the research field and in order to get a reasonable accordance in the distribution of citing documents over the time period (1974-1996) between the two selected articles. All in all some 160 records of citing documents were retrieved and saved. From these a total of 68 articles were acquired for analysis; 16 citing the one, 45 citing the other and 7 citing both of the originally selected articles.

3. The third step in the method was the analysis of cited and citing documents. A record for each document was created in a special database, comprising besides bibliographic data also all citations in context, classification of the citing document as for theoretical approach and of citation functions. When several citations to the same source appeared in the same citing document each citation was classified individually. At the same time, multiple classification was permitted, so that a single citation could often be seen as serving several functions simultaneously. The analysis and classification of functions centered naturally on the immediate context of the citation in the citing document, i.e. on the sentence in which the reference appeared and on those immediately preceding and succeeding it. Quite often, though, this was not enough to get a proper understanding of the function of the citation in the text. It sometimes proved necessary to look at the overall purpose of the citing document in order to see how the information provided by the citation "fit in" with the rest of the text. The analysis in itself involved three steps:

a) The first was the identification of possible instances of connectional motivations for citing (Vinkler 1987; cf. Liu 1993, p. 395). For this purpose, dependencies such as personal ties between cited and citing author, manifested e.g. through documented previous co-authorship or acknowledgement of indebtedness, or self-citations were registered. Only seven cases of documented personal ties were found (admittedly, the real number may be higher, since information about such ties is not always easily retrieved or even documented). The decision was made to regard any citation appearing in one of these documents as a suspect of "non-relevance", if its only function seemed to be that of providing background information. In none of these cases of documented personal ties, however, the citations could be dismissed out of hand as purely connectional, that is, as not relevant. The citations in question all proved either to serve entirely different functions than that of providing background information, such as serving as evidence, or served also at least one additional function, such as that of providing motivation for the research reported in the citing article, apart from that of giving background information. Consequently, all these cases were retained among the other citing documents for the next step in the analysis. Another seven documents in the sample contained self-citations. It was decided to keep these also for further analysis, to
find out if there was any marked difference in citation functions between these
documents and those of the rest of the sample.

b) The next step in the analysis was the classification of all documents as for theoretical
approach and citation functions, in accordance with what was stated above. It should be
noted that the categories used for classification of functions were developed in the main
directly from the material at hand, with only one eye to the typology of topical
relevance relationships that was developed by Green (1995b). After the initial analysis
was completed, the categories were more sharply delimited during a second tour. Only
then a comparison with other systems of classifying citation relationships was made,
which resulted in some of the observations made in section 3.2 about similarities
between categories used by different classification schemes. The categories used for our
case-study are presented and explained by means of examples in chapter five.

c) The third step in the analysis was the identification of topic matching relationships
between cited and citing documents, in accordance with the criterion specified in
section 1.1, i.e. having at least two descriptors in common in one of Dialog's scientific
databases. This part of the analysis could be performed only on part of the original
sample of documents, since in many instances there were no records for both cited and
citing documents in the same database. As a result, only 37 articles, the two cited
sources included, were analyzed for topic matching. The database selected for this
purpose was Aerospace database (Dialog file no. 108), since it proved to contain the
greatest number of records for the citing articles making up our sample. It was observed
that the two originally selected articles fulfilled the criterion of topic matching in this
database. For each of the remaining 35 articles having a record in Aerospace database
the number of shared descriptors with these two, cited and non-cited article alike, were
noted. This was made in order to find out if topic matching appeared to be more
frequent between cited and citing documents, than between documents between which
no citation relationship obtained. In addition to the mere counting of number of shared
descriptors, the so-called union overlap measure was applied to the same 35 documents
to express the topical relatedness between these and the two originally selected articles.
This measure, also known as Jaccard's coefficient or Hooper's equation, defines the
similarity between two documents with two sets of indexing terms, A and B, as:

\[ \text{sim}(A, B) = \frac{n(A \cap B)}{n(A \cup B)} \]

where \( n \) is the number of indexing terms (Harter 1993, p. 548). That is to say, the
similarity between the documents is measured as the quotient between the intersection
and the union of the sets of indexing terms of both documents. The range of the
resulting score is between 0 (complete dissimilarity), and 1 (complete similarity, i.e.
identical index terms). The weakness of this measure, though, is that it is vulnerable to
interindexer inconsistency.\(^{25}\) Harter (ibid.) reports on studies in which documents that

\(^{25}\) Another major weakness of the union overlap measure is that it takes no account of
the differences in discriminating power between index terms. Thus, e.g. two documents
that have both been indexed by only one, very broad index term, that happens to be the
were in fact identical, hence, ought to have a score of 1, have shown similarity scores as low as 0.5, due to interindexer inconsistency. In the study of three main journals in library and information science undertaken by Harter and his colleagues the mean similarity score for citing and cited articles, as measured with the indexing terms taken from ERIC, LISA and *Library Literature*, ranges from a low 0.096 to high 0.391, between citing and cited documents. Keeping this in mind and employing in accordance with *Harter* (ibid.), a strict all-or-nothing attitude,26 the computed similarity score of the same for both documents, would get a similarity score of 1 - complete similarity, no matter what how many other documents have been indexed by that same term in the database. In fact, in the case of *London* (1974) and *Molina* (1974), they are both indexed in EMBASE by the one descriptor "environmental health". They will thus have a similarity score of 1 in EMBASE (provided only descriptors are used for the computation, of course). However, a search performed with "environmental health" as sole query term, limited to the descriptor field alone, will retrieve no less than 13000 items (1 August 1996). On the other hand, a search query with the two descriptors shared by *London* (1974) and *Molina* (1974) in Aerospace database, viz "ozone" and "atmospheric chemistry", retrieved a considerably lower number, 1932 items in that database (1 Aug. 1996). Yet, due to the higher number of descriptors assigned to each document in Aerospace database, the similarity score for *London* (1974) and *Molina* (1974) is only 0.1. This seems to run counter to the intuition, that documents sharing a property that is shared by relatively few other documents are somehow "closer" or more "similar", than documents that have only some very general property in common, a property which is shared by a very large number of other documents. What seems to be needed is a measure of document "similarity" or "relatedness" that is not biased towards indexing systems that assign relatively few, broad index terms to documents. Such a measure should ideally reflect also the dynamical character of the "relatedness between documents, since as information grows, the "docuverse" expands, the "distance" between documents will sometimes tend to grow too.

26 This approach means simply that anything less than a perfect character-by-character match will count as no match at all, thus leaving no room for partial matches between such indexing terms as terms as e.g. "atmospheric chemistry" and "atmospheric composition". This is an easy-to-use, unambiguous method of judging about similarity between documents. The disadvantage with this way of measuring, as noted by both *Harter* (1993, p. 549) and *Green* (1995b, p. 649), is that it will tend to underestimate, sometimes grossly, the topical relatedness between two documents. However, in the present study this all-or-nothing approach can be justified, I think, by the fact that it is applied only to descriptors, that are part of a controlled vocabulary, in which terms most often have a specific, limited meaning. Furthermore, since the database used for the computing of the *similarity score* specializes in a limited subject area, obviously the criteria for what will qualify as "topical relatedness" within the database will be rather strict. Thus, for example, "atmospheric chemistry" and "atmospheric composition", both descriptors in the Aerospace database, evidently stand for different concepts in this database, since there are records, notably *London* (1974), that have been indexed by both these terms.
0.1 for the two originally selected articles, as measured with descriptors from Aerospace taken as the indexing terms, may not be considered so low as it appears at first sight. In fact, depending on the total number of indexing terms assigned to the documents, it is quite conceivable to find instances of "topic matching", in accordance with the criterion specified above, that would show an even lower similarity score. Thus, the union overlap measure produces no absolute value of the similarity between two documents; it is dependent on such things as databases selected, indexation depth, indexing terms used, number of index terms assigned etc. All this said, the union overlap measure has still been considered useful in the present study as one of several tools for determining the relationship between documents.

Not being in a position to rule out a topic matching relationship in all those cases where records for both cited and citing articles were not found in the same database, the claim is made, rather, for the diverse, multi-purpose use of citations, possibly reflecting also the multi-valued character of the concept of relevance. At best, then, the citation analysis performed here will provide empirical evidence for the claim that there are at least some cases in which topic matching is neither sufficient nor necessary for describing relevance relationships.
5 A case study: monitoring of the ozone layer

5.1 Background: stratospheric ozone monitoring

Ozone - O₃, is a gas existing both on earth near the ground and in the atmosphere. The former is poisonous and has an adverse effect on plants in our immediate environment. The atmospheric ozone, on the other hand, plays a double rôle: it has a negative influence on human environment in the sense that it may contribute to the so-called greenhouse-effect, the warming up of the earth's climate that threatens, among other things, to raise the sea level so that areas that are now land will be completely flooded. But, at the same time the atmospheric ozone is also of vital importance for life on earth, since it serves as a protective shield, preventing too much biologically harmful solar ultraviolet (UV) radiation from reaching the earth's surface. A reduction of the atmospheric ozone may lead to a global increase of skin cancer incidents, and also have other adverse human health effects (Huwyler 1992).

The atmosphere is commonly subdivided into several layers. Nearest to the Earth is the troposphere, which is the region where clouds form and which contains the largest percentage of the total air mass (about 80%). It stretches out between 8 km (above the poles) to 18 km (above the equator) from the Earth's surface. The tropopause, with an average height of 11 kilometers, is the intermediate layer, dividing the troposphere from the stratosphere. The stratosphere ranges in altitude between 10-50 km above ground level. At still higher altitudes follow the mesosphere the thermosphere and the exosphere. The protective ozone layer is located at an altitude between 20-30 km in the stratosphere. Hence, we speak of it as stratospheric ozone.

Ozone is constantly produced in the stratosphere through the splitting up of oxygen molecules -O₂, by means of UV-radiation. The single oxygen atoms that result then combine with other oxygen molecules to form ozone - O₃. At the same time, stratospheric ozone is also constantly broken down and destroyed by another type of UV-radiation. In this way a balance is upheld, involving also the transportation of air, especially from the equator (cf. Nolin 1995, p. 56). However, since ozone is a highly reactive gas, which tends easily to "give away" its third surplus atom to become again stable oxygen O₂, the balance can be disturbed by external influences from substances that will react with the ozone. But, on the other hand, it should also be kept in mind that, although there is a natural balance, the amount of ozone in the stratosphere varies naturally from time to time, e.g. as a result of variations in solar UV radiation with sunspot cycle (Bauer 1979). These natural changes make it more difficult to detect long-term trends that are caused by other forces, such as the emission of certain man-made chemical compounds.

---

27 For an excellent overview of the atmospheric structure and the function of ozone in it, adapted to a lay reader, consult the WWW-site produced by NASA Goddard, at URL: http://daac.gsfc.nasa.gov/CAMPAIGN_DOCS/ATM_CHEM/ozone_atmosphere.html
Nolin (1995) distinguishes between three separate areas of research concerned with stratospheric ozone, viz. "monitoring of the ozone layer, research on the effects of a decreased ozone shield and industrially sponsored research on alternatives to, or substitutes for, those compounds that have been considered too dangerous for use" (op.cit., p. 244; my emphasis). The process of "functional integration" (ibid.) between these three areas are dealt with only incidentally in this thesis. In accordance with Nolin (1995), the focus of the case-study is on monitoring. This is a field of research in which the scientists involved try to measure the amount of stratospheric ozone with different methods, study the changes, make predictions of future development and draw conclusions about the processes behind the observed measurement data. Researchers from several disciplines are working in the field; main actors come from disciplines such as chemistry, meteorology and physics. Owing partly to the heterogeneous composition of the field's scientific community, two principally different theoretical approaches can be discerned: the "chemical" and the "dynamical". The "dynamically" inclined researchers generally seem to have a stronger meteorological orientation, counting primarily with the movements of the circulating air and other natural causes as affecting the ozone layer more than anthropogenic causes, such as the emission of chemical substances from various polluting sources (op. cit., p. 60). The "chemically" minded researchers, on the other hand, tend to disregard the effects of the circulation of air when computing the chemical reactions affecting the stratospheric ozone, since taking these movements into account would make their chemical and mathematical models very complicated (ibid.). In the case study performed here one of the points of the analysis was to examine if the articles under consideration could be classified as representing either theoretical approach, the "dynamical" or the "chemical", in order to find out to what extent and in what way this factor seemed to be correlated with citation behavior. For this purpose, special care was taken to base judgment whether or not to classify an article as either "chemically" or "dynamically" oriented exclusively on the contents of the article under consideration, disregarding the authors' background and affiliation or the references made. This is an important reason why in several cases it was not possible to arrive at a clear cut decision about how to classify the theoretical approach of an article, as either "chemical" or "dynamical". (In fact, in at least one case the same scientist appears as the author of both a "chemically" and a "dynamically" oriented article.) The grounds for classifying an article as either "chemically" or "dynamically" oriented (henceforth "C" and "D" respectively) were the following:

The articles classified as C are restricted to discussing the effects of various chemical substances on the ozone layer. They further attribute the emission of these various chemical substances in the main to anthropogenic causes. A typical example is Bruckman (1988), an article that is concerned with the concentration of chemical trace compounds in urban air, singling out traffic and industry as the main sources of emission:

In addition, the impacts of the accumulation of halogenated compounds in the environment are discussed, along with the unfavorable changes caused by the penetration into the stratosphere of those halogenated hydrocarbons with lifetimes of several decades in the atmosphere. (op. cit., p. 2364)
The second requirement for classifying an article as C, that the chemical compounds affecting the ozone layer be attributed to anthropogenic causes, is important, since there are some articles which clearly discuss effects of chemical substances, but which tend to look for natural sources of their emission. The following excerpt may serve as a typical example:

What then is the source of CCl₄? Marine algae emit halocarbons (5) and a biological source of chlorocarbons is possible but its strength is unknown. A natural, probably atmospheric, origin of CCl₄ must also be considered...

(Lovelock 1974, p. 293)

Articles of this latter inclination, then, have been classified as D, as have also those articles which discuss mainly the effects of air circulation, solar proton events (i.e. the emission from time to time of large fluxes of protons, together with smaller quantities of helium and other heavy atomic nuclei from the sun), and other natural causes affecting the stratospheric ozone, while tending to play down the importance of the influence of chemical agents. To this category belong also articles that point specifically to the self-healing, balancing forces inherent in nature or that rely heavily on measurement data and assert a natural, cyclic variation in the levels of ozone measured.

The following extract is an example representing this latter type of article:

The long-term nonlinear trend in the water vapor data showing an increase in the 1960s and a decrease in the 1970s indicates that linear trends deduced from a few years of data are only applicable to that data period, and the use of such trends to infer future trends is highly questionable. The long-term trend in stratospheric water and its similarity to the long-term trend in stratospheric ozone suggest that these changes arise from long-term changes in the intensity of the circulation.

(Mastenbrook 1983, p. 2164)

Thus, the principal beliefs that constitute the two different theoretical approaches, or "metaphysical paradigms", seem to be grounded in two different models of our planetary environment: either as a fragile entity, that is subject to "external", human influences, and that may in fact have its natural equilibrium disturbed to such a degree, that life on Earth is threatened (C). Or, alternatively, it is seen as essentially robust and sturdy, while constantly and naturally changing, possessing self-healing powers that allow for adaptation to new conditions (D). This is made explicit in an article by Hare (1980), symptomatically entitled "The planetary environment : fragile or sturdy", in which the author talks about "the natural resilience of the world" and further makes the confession that

I suspect that it can absorb more insults than we suppose. So far, for example, we cannot detect in the climatic record any remarkable impacts due to human action.

(op. cit., p. 363)

Articles that have been considered impossible to classify as belonging to either category (C or D), are those that either make no clear statement as to the causes affecting the changes in stratospheric ozone, or that seem to devote equal attention to anthropogenic and natural causes. Typical instances of this kind are review articles, the purpose of
which is to sum up the scientific discussion in the field and to expound the arguments put forward by both camps. In general, a policy of cautiousness has been applied, so that in any case of even the slightest hesitation an article has been left unspecified (U), rather than forcing a theoretical approach upon it. This means, admittedly, that the figures for C and D in Tables 2 and 3 (below section 5.4) may in fact be underestimated, but supposedly not to an extent that would alter the overall tendency shown by the results.

5.2 The material

The articles selected as point of departure for the citation analysis were both published in 1974; one in *Nature*, the other in *Science* - two well-known, prestigious, multidisciplinary scientific journals. Garfield (1979), in particular, as a result of a citation analysis conducted in 1977 of the Journal of Geophysical Research (JGR) - the most frequently cited journal in the geological and geophysical literature, has assessed the role of these two journals:

The importance of *Science* and *Nature* in the research continuum that stretches from geology to astrogeophysics was demonstrated by the fact that both of them ranked among the top five journals cited by JGR and among the top seven sources of citations to JGR. The latter point shows the strong geophysical orientation of the two multidisciplinary journals. Undoubtedly, the review quality of many of the articles published in Science and Nature account, to a significant extent, for the frequency with which JGR cited them. (op.cit., p. 186)

The year 1974 was selected, since it constitutes something of a turning point in the development of stratospheric ozone monitoring. It can be described as the beginning of a new phase in the controversies occurring in this field of research (Nolin 1975, p.100). An earlier controversy, in the late 1960s and beginning of the 1970's had mainly concerned the potential effects on the ozone layer of a future fleet of supersonic aircrafts, the so-called SST (Super Sonic Transport). Scientists opposed to the SST pointed first to the threat of emission of large quantities of water vapour, which would react with the atmospheric ozone and destroy large parts of it, thereby causing thousands of new cases of skin cancer each year. The water vapour theory seems to have been discarded after a scientific conference in 1971, but was replaced by a second theory, according to which emission of nitrogen oxides, NOx, from the SST would lead

---

28 Failing to compute the *impact factor* (the average number of citations per published article) for the year 1974 (in order to do that it would have been necessary to find out the total number of citations to articles published by the respective journals that year), it was nevertheless observed, that the total number of documents that cite the two journals during the time period 1974-1996, as recorded in the SciSearch database, is on the same level; 1290752 for Science and 1565403 for Nature. This means that the rate of citedness is somewhat higher for Science, since apparently it publishes fewer articles (at least there are fewer records of articles appearing in Science in the database.)
to a serious depletion of the stratospheric ozone (op.cit., pp. 92ff). The dispute over the SSTs was not yet settled, when in 1974 two scientists, Sherwood Rowland and Mario Molina, from the Department of Chemistry, University of California at Irvine, introduced a theory to the effect that chlorofluorocarbons, CFCs, by that time widely used as propellants in aerosol sprays and as a cooling liquid in refrigerators, were posing an even more serious threat to the earth's ozone shield than the SST. The catalytic chain reaction between the chlorine atoms and ozone was computed by Molina and Rowland to be "considerably more efficient" (Molina 1974) than the corresponding reaction between nitrogen and ozone in the destruction of atmospheric ozone.

The basic research results behind the CFC-theory were first published in a short paper in Nature, June 1974. This article, entitled "Stratospheric sink for chlorofluoromethanes: chlorine atom-catalysed destruction of ozone", 29 is thus one of the two documents selected as cited sources for the citation analysis in this case study. Henceforth it is referred to as Molina (1974). However, the CFC-theory did not draw wide attention until September that year, when, following a pressconference given by Molina and Rowland at the conference of the American Chemical Society in Atlantic City, the news of a possible CFC-caused 30-50% reduction of the stratospheric ozone to the year 2050 was on the cover of the New York Times (Rowland 1987; Nolin 1995, p. 103). Although the reaction from both the media and from the scientific community up until then had been rather lame, the CFC-theory had already received support from a few other scientists, notably from Richard Stolarski and Ralph Cicerone, who had themselves discovered the chlorine-initiated chain reaction destroying the ozone the year before, but had not yet published their results, when Molina and Rowland independently thereof made the same discovery (Rowland 1987). It is not surprising, then, that the article eventually published by Stolarski and Cicerone (it appeared in the special issue dealing with the ozone problem of the Canadian Journal of Chemistry, volume 52, number 8) is frequently co-cited with Molina (1974). In fact, there is already a reference in Molina (1974) to Stolarski's and Cicerone's article, then still in press. According to Nolin (1995, p. 103), Molina and Rowland soon after publication entered into contact with Stolarski and Cicerone, and it was decided that the latter two would publish a presscommunique in support of the CFC-theory.

Two other scientists that seem to have played important roles in the formation and early reception of the CFC-theory were Harold Johnston and Paul Crutzen, interestingly enough both meteorologists, although with a clear interest in atmospheric chemistry. Both these scientists had been involved already in the debate over the SST; they were the main proponents of the second SST-theory, according to which nitrogen oxides from the SST, through photolytic dissociation in the upper atmosphere, would enter into

29 The title of the article was actually misprinted when it was published, so that the subtitle read: "chlorine atom-catalysed destruction of ozone". Although bibliographically incorrect, the title given here was chosen, for the purpose of clarity, and since it is also the title given by one of the authors in a reference to the same article (cf. Rowland 1975). Other variations of the title have also been proposed, notably in SciSearch.
a catalytic chain reaction with ozone that would eventually lead to a serious depletion of the ozone layer. According to one computation made by Johnston, a fleet of 500 SST would possibly lead to a reduction of stratospheric ozone of 10 percent in only two years (cf. Nolin 1995, p. 95). Crutzen had already in 1970 observed the influence of NOx on atmospheric ozone, and in particular the possible threat from the SSTs. Even more important for the purpose of this case study, he had performed model computations of ozone reduction caused by CFCs. Articles of both Crutzen and Johnston are cited by Molina (1974), and there is also an explicit acknowledgement of a helpful discussion with the latter of the two researchers. Nolin (1995, p. 103) further points to an "important exchange of views" that apparently shall have taken place between Crutzen and Rowland soon after publication of Molina (1974).

The basic idea of Molina (1974) is perhaps best expounded by its short abstract:

Chlorofluoromethanes are being added to the environment in steadily increasing amounts. These compounds are chemically inert and may remain in the atmosphere for 40-150 years, and concentrations can be expected to reach 10 to 30 times present levels. Photodissociation of the chlorofluoromethanes in the stratosphere produces significant amounts of chlorine atoms, and leads to the destruction of atmospheric ozone.

(Molina 1974, p. 810)

The main point, thus, is that although the amounts of chlorofluoromethanes or CFCs that are released into the atmosphere may appear to be small in comparison with that of other pollutants (the figures reported in the article represent the world production rates for 1972; they were 0.3 and 0.5 Mton per year respectively for the two chlorofluorocarbons CF2Cl2 and CFC13, but were then still increasing), due to the chemical inertness of the CFCs, and since no other sink has been found, i.e. no other substance reacting with these compounds on their way up towards the stratosphere, the effect on the stratospheric ozone will be dramatic. ("Petite cause, grand effet", as Blaise Pascal once had it.) The article does not, however, speculate so much over the potential consequences, although it is obvious that the authors are worried:

It seems quite clear that the atmosphere has only a finite capacity for absorbing Cl atoms produced in the stratosphere, and that important consequences may result. This capacity is probably not sufficient in steady state even for the present rate of introduction of chlorofluoromethanes. More accurate estimates of this absorptive capacity need to be made in the immediate future in order to ascertain the levels of possible onset of environmental problems.

(Molina 1974, p. 812)

Molina (1974) must clearly be classified as a C-article, since it deals almost exclusively with anthropogenic causes affecting the ozone layer. It is explicitly pointed out in the
article that the CFCs "do not occur naturally (except possibly in minute quantities from volcanic eruptions)" and that "rapid biological removal seems unlikely" (ibid.).

The other article selected as one of the cited sources apparently has had a weaker impact on the scientific community within the research field than Molina (1974). It was published in Science, May 1974, and written by two researchers from the Department of Astro-Geophysics, University of Colorado at Boulder, Julius London and Jean Kelley. This article, entitled "Global Trends in Total Atmospheric Ozone", henceforth referred to as London (1974), presents an analysis of data from measurements of global distribution of ozone for the 13-year period from 1957 through 1970. For this period, the analysis arrives at an upward trend, i.e. an increase in total ozone of about 7.5 percent per decade in the Northern Hemisphere and about 2.5 percent per decade in the Southern Hemisphere. No definite conclusions as to the possible cause of these trends are drawn. However, although the article explicitly mentions the second SST-theory, making reference to Johnston (1971), there seems to be a clear inclination towards "dynamical" explanations:

We have yet to establish whether these changes show a definitive worldwide pattern and to determine in what manner these trends are associated with changes in solar radiation, general stratospheric photochemistry, or fluctuation in the stratospheric circulation. The marked hemispheric difference in the trends, however, would seem to preclude any direct solar influence.

(London 1974, p. 989)

The last remark in the quoted passage is interesting, since it points to a possible source of controversy, as for the interpretation of data. Other, more "chemically" oriented scientists have interpreted this marked difference in trends between the two hemispheres as a clear indication of the impact on stratospheric ozone of the atmospheric nuclear bomb tests that were performed in the Northern Hemisphere in the beginning of the 1960's. In particular, this was the interpretation made by Crutzen (1975) and Rowland (1975), among others. The authors of London (1974) would perhaps not be opposed to such an interpretation, but obviously they have chosen a more restrictive, perhaps more prudent attitude.

---

31 It is, by the way, interesting to note the different uses Molina (1974) and London (1974) make of Johnston (1971). For the authors of the former paper, Johnston's article, together with other papers - among them Crutzen (1971), serves mainly as an object of comparison and a piece of evidence supporting their own claims about the influence of chlorine on ozone. For London (1974), the only function of the reference to Johnston (1971) seems to be, rather, to emphasize further the general importance of monitoring atmospheric ozone, motivating thereby their own research. We will see more examples of the same kind in the case-study, when we look at how differently the two articles, London (1974) and Molina (1974) are used.

32 For a discussion of the impact on atmospheric ozone of the nuclear tests of the 1950's and 1960's, see also Chang (1979)
When comparing London (1974) and Molina (1974) it is important to note first the difference in time perspective. Although both articles were published in the same year, within a months time span, the orientation of the two is very different; the former being oriented towards the past, the other towards the future. London (1974) could, of course, be used for making inferences, by means of extrapolation, about future trends as well (and in fact it has been used that way too, at least in the beginning of the controversy over the possible impact of the CFCs on the stratospheric ozone). But there is nothing in the article itself that would seem to either encourage or discourage from such use. By contrast, Molina (1974) makes explicit reference to the future, when it comes to evaluating the consequences of the CFC-theory:

One important corollary of these calculations is that the full impact of the photodissociation of CF2Cl2 and CFC13 is not immediately felt after their introduction at ground level because of the delay required for upward diffusion up to and above 25 km. ... A lengthy period (of the order of calculated atmospheric lifetimes) may thus be required for natural moderation, even if the amount of Cl introduced into the atmosphere is reduced in the future. (op. cit., p.812)

Obviously this means, among other things, that the data about ozone trends presented by London (1974), cannot possibly be used as evidence, at least not in a strict logical sense of the word, either against or in support of Molina (1974) and the CFC-theory. An evaluation of the CFC-theory, in the version first expounded in Molina (1974), against empirical data will be possible only in a rather distant future.33

How, then, should the topical relationship between the two articles be described? In particular, can a relationship of topic matching in accordance with the definition given above, be said to hold between them? Looking at the indexing made in a few international scientific databases we observe that London (1974) and Molina (1974) are both indexed in EMBASE (Excerpta Medica, produced by Elsevier Science; Dialog file no. 73) and Aerospace database (Dialog file no. 108). The former of these two databases is specialized in "human medicine and related disciplines" (Dialog, 1995), and its subject coverage includes "environmental health". Both London (1974) and Molina (1974) are indexed in EMBASE with the general descriptor "*environmental health (0015525)". However, since this database has went through major changes in its indexing policy since the publication of these two articles, it was found unfit for the purpose of making comparisons with the indexing of citing articles.34

33 A possible example of the different uses made of the two articles due to their differing time-perspectives is provided by Heath (1982), which cites London (1974) but not Molina (1974), whereas in Heath (1988), which deals with the period 1970-1986, the roles are reversed, i.e. Molina (1974) is cited but not London (1974).
34 E.g., for records added to the database prior to 1988 it seems only broad, general descriptors were assigned; instead, a deeper indexation was given by EMCLAS Section Headings.
The other database in which records for both London (1974) and Molina (1974) were found, the Aerospace database, is described as supporting basic and applied research in aeronautics, astronautics, and space sciences, as well as technology development and applications in complementary and supporting fields such as chemistry, geosciences, physics, communications, and electronics.

(Dialog, 1996)

Clearly, this database is "closer" to the disciplines, at the crossroads of which is situated the research field concerned with monitoring of the stratospheric ozone layer. This fact is also reflected through the descriptors assigned in Aerospace database to the two articles under consideration. Both London (1974) and Molina (1974) are indexed by the descriptors "atmospheric chemistry" and "ozone". On top of that, each article is indexed with an additional nine descriptors, that are not shared with the other. From these it would seem fairly obvious to someone who had not yet read either article that London (1974) is concerned with measurement, whereas Molina (1974) deals rather with chemical reactions. This impression is further strengthened when it comes to subject classification; London (1974) is classified by Aerospace database under "6513 Geophysics (1965-74)", while Molina (1974) is placed under "6506 Chemistry (1965-74)".

Apart from the two databases mentioned above, in which records for both the selected articles were found, yet another record for London (1974) was retrieved in Pascal (Dialog file no. 144), a bilingual French-English multidisciplinary database produced by the Institut de l'Information Scientifique of the French National Research Council. The descriptors assigned by this database to London (1974) were also retained for comparison, when applicable, with those of citing articles.

We already noted that the two articles selected do fulfil the criterion specified for topic matching. It must be remembered, however, that a topic matching relationship between two documents is never absolute, but something that holds only with respect to a specific search query, under specific conditions of inquiry and from a specific viewpoint. Given that someone, searching the Aerospace database, would be interested in finding literature about atmospheric ozone, it would seem that both London (1974) and Molina (1974) could possibly offer that person some valuable information. This is not to say, however, that they would offer the most "relevant" information, the information that would best satisfy that persons need to know. It must also be remembered, that one of the points of making a "subject search", may be precisely to find documents that treat the same subject or topic from different angles or viewpoints, using different methods or arriving at different conclusions. If all the articles retrieved would contain exactly the same information, if there would be no differences whatsoever in either theory or method applied or conclusions drawn, we would surely have a case of 100 percent "topic matching", but at the same time - 100 percent redundancy!
There are, to be sure, other measures of relatedness between two documents, than the topic matching criterion used here. One such measure is that of bibliographic coupling, that was introduced in the 1960's by M.M. Kessler of the Massachussets Institute of Technology. Bibliographic coupling is the number of shared references between two papers; the value of the relationship is said to be of "strength one, two, etc. depending on the number of shared references" (Martyn 1964). According to this measure, the relationship between London (1974) and Molina (1974) is of strength one, since they have one reference in common, a paper on SST by the above mentioned Harold Johnston entitled "Reduction of stratospheric ozone by nitrogen oxide catalysts from Supersonic Transport exhaust", which was published in Science in 1971. But the weakness with bibliographic coupling is apparent already from this single example. As has been pointed out by Martyn (1964) and Weinberg (1994), the fact that two articles have a reference in common does not necessarily mean that the citation to that reference is to the same part of the document. It may well be that the two articles cite completely different points of the same document, that the citation is made for completely different reasons in the two cases. It is one of the points of this thesis, to show how different citations to one and the same document serve different functions and that the citing documents are sometimes also related in different ways to the topic of the cited document. In the case of the one reference shared by London (1974) and Molina (1974) the topic may well be close in both citations, but the functions are clearly different. In London (1974), the citation to Johnston (1971) is more general and serves primarily as part of a motivation (cf. below) for engaging in ozone research at all:

Added interest in studies of atmospheric ozone stems from recent suggestions that exhaust products from high-flying aircraft might somehow affect the stratospheric ozone distribution and thus change the ultraviolet shielding property of the atmosphere, or inadvertently modify the earth's climate (5).

(London 1974, p. 987; ref 5 is to Johnston 1971)

Note that this citation actually does not endorse the information contained in the cited document, but appears to remain neutral towards it. From the rest of the article it seems clear also that the authors do not in any way make use of this information, in order to develop their own argument. By contrast, in Molina (1974) the function of the citation to Johnston (1971), in conjunction with citations to other documents, is to explain, by means of a comparison, the role of chlorine in the destruction of atmospheric ozone:

The odd chlorine chain (Cl, ClO) can be compared with the odd nitrogen chain (NO, NO2) which is believed to be intimately involved in the regulation of the present level of O3 in the atmosphere (7-10).

(Molina 1974, p. 810; ref. 8 is to Johnston 1971)

Another method of capturing the similarity between two documents is to measure term co-occurrence, for example, in titles and abstracts. This method has been used by Donna Trivison (1987; see: Harter 1993, p. 545). She found that although term co-occurrence was in fact much greater for cited/citing document pairs than for randomly selected pairs of documents, nevertheless as much as 23% of the cited/citing document pairs were not "term-related", in the sense of sharing at least five significant word stems in
titles and abstracts. Applying the same method on *London* (1974) and *Molina* (1974) it is clear that they also are not "term-related" in this sense; the only significant term they have in common in both titles and abstracts is, not surprisingly, "ozone".

The sample of articles studied, *London* (1974) and *Molina* (1974) excepted, consists of sixty-eight articles, the overwhelming part of them from scientific journals. Each article in the sample cites either *London* (1974) or *Molina* (1974) or both, at least once. No claim is made as for the representativity of the sample. Although the selection was fairly random, albeit not in a strictly scientific sense of the word, it was also influenced to a certain extent by the accessibility of the respective journals in which the citing articles appeared. Also, the proportion in the sample between articles citing *London* (1974) and *Molina* (1974) respectively does not reflect the actual rate of citedness for the two articles. A search performed in SciSearch (Dialog file no. 434) in January 1996 resulted in 898 citations for *Molina* (1974), but only 55 for *London* (1974). Several possible explanations can be adduced for the significantly higher citation rate of *Molina* (1974). One likely such explanation may be the fact that, as already noted, whereas *London* (1974) is oriented essentially towards the past, *Molina* (1974) is more future-oriented, thereby, as it were, inviting to further research. And since further research into a matter often builds on earlier research, a possible side-effect thereof might be, precisely, a higher rate of citations to the "inviting" article. In general, it could be suspected that documents in which hypotheses are put forward, that can be either substantiated or refuted by future research, will draw a higher number of citations, than documents that refrain from making any specific statements about the possible causes of natural phenomena, but content themselves with the presentation and interpretation of measurement data. Another more obvious reason for the marked difference in citation rate between the two articles may be the fact that *Molina* (1974), so to speak, represents the "winning paradigm", whereas *London* (1974), although not explicitly denying the influence of anthropogenic, chemical sources on the stratospheric ozone layer, appears to be more associated with a "conservative" arriere-guard in the field of research under consideration. Whereas *Molina* (1974) introduces novelty in the field, thereby inciting new stakes to be made, *London* (1974) is more concerned with defending positions already upheld. And, of course, when placing ones bets, most people prefer to place them on the winner. This assumption seems to gain further support if we look at the distribution over time of the sources citing the two articles respectively (Figure 3). As many as thirteen, almost a quarter of the total number of citations to *London* (1974) in SciSearch come from articles published in 1975. During the period 1976-1983 it has a yearly rate of citedness in the order between two and five. Then follows a period of four years without a single citation, after which it receives only sporadic citations in 1988 (2), 1992 (2) and 1994 (1). By contrast, *Molina* (1974) is cited 3 times during the year of its publication and the following year it already receives an impressing 66 citations, followed by a further increase to 79 in 1976. An expected decrease during the following years does occur, but the rate of citedness is still kept at a high level and as yet has never, except for 1974, been below 20 citations (1982) in one single year. During recent years, there seems even to be an increased interest for *Molina* (1974), since its rate of citedness has again risen; for 1995, more than 20 years after its publication, it received according to SciSearch no less than 71 citations. In short, then, it seems that as time
passes, and evidence apparently in support of the CFC-theory accrues, the difference in
citation rate between the two articles becomes more pronounced. This appears to
continue up until today, 22 years after the publication of the two articles. A new search
in Scisearch a couple of months later, in March 1996, revealed that Molina (1974) had
gained another 10 citations, whereas London (1974) had not advanced one single step.
This all is hardly surprising. The phenomenon is well-known by information scientists.
It has been called the "Matthew effect" in science (cf. Garfield 1979, p. 255), and it
means simply that success often breeds success - in science, as in life in general.
Citation counts may even tend to enhance this effect. As Martyn (1975) observes,
"common sense suggests that the most heavily cited materials are those which have a
higher probability of being cited in future" (op. cit., p. 291).

Figure 3: Citation rates for London (1974) and Molina (1974) during the time period
1974-1995, according to SciSearch (Dialog file no. 434)

The selection process involved several steps. First there was an initial search in
bibliographical records for articles citing London (1974) were saved and printed out.
For articles citing Molina (1974) a further selection was made. The search was limited
to specific years of publication, with consideration to the development within the field
of research studied, while at the same time an effort was made to arrive at a reasonable
distribution of articles over the time span covered. For example, the publishing years
1987 and 1988 were selected as important, being the year of international agreements (on a 50% reduction of CFCs until 1999) manifested through the Montreal-protocol (September, 1987) and the year of the large conference in Snowmass, Colorado (May, 1988), which appears to have been crucial in building consensus and settling the dispute between "chemists" and "dynamists" over the Antarctic "ozone hole" (cf. Nolin 1995, p. 130). On the other hand, a closeness in time was sought between the publishing years of articles citing London (1974) and of those citing Molina (1974) that were selected for analysis. As a result, the records already retrieved of references to London (1974), influenced to a certain extent the choice of publication years for articles citing Molina (1974). However, it was not always possible to attain such closeness in time. For example, only two articles published 1987 or 1988 that cite London (1974) are recorded in SciSearch. One of these is included in the sample of articles analyzed. All in all, 105 bibliographical records for articles citing Molina (1974) were retained and printed out; of these 52 articles were later acquired for analysis. Out of a total of 55 articles citing London (1974), 23 were acquired and analyzed in the case study. According to SciSearch (Jan. 1996), all in all there were only 13 articles that cited both London (1974) and Molina (1974); of these 7 were acquired for analysis.

<table>
<thead>
<tr>
<th>Py</th>
<th>L</th>
<th>M</th>
<th>L&amp;M</th>
</tr>
</thead>
<tbody>
<tr>
<td>1974</td>
<td>-</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>1975</td>
<td>3</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>1976</td>
<td>1</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>1977</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>1978</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1979</td>
<td>3</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>1980</td>
<td>1</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>1981</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1982</td>
<td>2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1983</td>
<td>3</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>1987</td>
<td>-</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>1988</td>
<td>1</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>1989</td>
<td>-</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>1992</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>1995</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>1996</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Σ</td>
<td>16</td>
<td>45</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 1: Distribution of citing articles in the sample over publishing years

From this already it is evident that there is a clear overrepresentation of articles citing London (1974) in the material of the case study, as compared with the total number of citations received by London (1974) and Molina (1974) respectively. Still, although no claims are made for representativity, the number and distribution over time of the articles citing Molina (1974) that were analyzed should be sufficient for giving an idea about different types of use that has been made of this article. The distribution of the
citing articles over publication years (Py) is shown in Table 1, where L stands for articles citing London (1974), M for those citing Molina (1974) and L&M for those citing both London (1974) and Molina (1974). (To be sure, the articles represented by the figures of the last column are not included in the figures of the two preceding columns.)

The median publishing year for articles of the second column is 1980.5, for those of the third it is 1980, and for those of the last column it is 1977. Seven of the articles in the sample that cite Molina (1974) are written by F.S. Rowland, co-author of Molina (1974). These articles, three of which are equally co-authored with M.J. Molina, have been included in the study in order to investigate if there appears to be any marked differences in function between self-citations and citations made by others.

Thus, the material of the case-study consists of 70 articles: the two selected as cited sources - London (1974) and Molina (1974), and the 68 citing articles. Each article, as already mentioned, was classified with respect to theoretical approach (T) into either of three classes:

\[ C=\text{chemical}, \quad D=\text{dynamical}, \quad U=\text{unspecified}. \]

Further, each citation to either London (1974) or Molina (1974) was analyzed in its context and classified according to its apparent function(s) (F) in the citing article. The classification scheme and the principles guiding its application is introduced, by means of examples from the material, in the next section.

5.3 Classification of citation functions

The categories used for the classification of citation functions in the case-study were the following:

\[ \text{ba=background, bi=bibliography, co=comparison, ev=evidence, ex=example, fo=force, ma=magnitude, me=method, md=modification, mo=motivation, ob=objection, qu=quote, re=result} \]

The categories are described below, with examples intended to illustrate their use:

**Background (ba):** Cited information reports on previous research, as part of the framework for own research, but is not being used in inference to prove or develop own position. This is the by far most common function of the citations in the material analyzed. It may sometimes appear to represent a "weaker" link between the citing and the cited documents, than most of the other categories. That is to say, that passages cited as background information could more easily be imagined as left out, without detracting seriously from the strength of the argumentation in the citing article.

---

35 One of these articles also has a reference to London (1974).
Background information typically appears in the introductory part of citing articles. As two examples, consider the following passages:

1. INTRODUCTION
Trends and fluctuations in monthly mean values of total ozone have frequently been noted (e.g. Komhyr et al. 1971; London and Kelley 1974; Kulkarni 1976; Angell et al. 1985); sometimes after a longer record is available it becomes apparent that a trend is just a fluctuation on a longer time scale. (Newell 1988, p. 595)

INTRODUCTION
Following Crutzen's (1970) and Johnston's (1971) identification of the threat to stratospheric ozone from catalytic destruction by increased levels of NO and Molina and Rowland's (1974) identification of the similar threat from Cl, major theoretical and experimental efforts have been directed toward a more detailed understanding of the ozone destruction problem and its relation to man's activities. (Goldan 1980, p. 413)

Bibliography (bi): Describing or tracing the bibliographical-literary history of cited work. This is a kind of meta-citation, for which it is hard to find a proper name. There is reason to believe it is more common in the humanities than in science. What is meant is a citation made not out of a primary interest in the contents of the cited work, but which seems to be motivated rather by a wish to provide background information for its interpretation, or demonstrate its significance, importance or impact on others. All four instances of this category in the material proved to be self-citations. This may not be surprising, since the strongest interest in asserting the impact of a particular article obviously lies with its authors. As a probably non-typical, but all the more telling example, consider the following extract:

Our initial short paper on the chlorofluoromethane-ozone problem was written and submitted to Nature in late January. (2) However, the scientific, political, and regulatory ramifications of this research were so widespread that we felt obligated to publish a much more detailed exposition than that contained in the limited space provided for the Nature paper. ... The work reported in this article and in its predecessor in Nature was the basis for the award of the 1983 Tyler Prize in Environmental Achievement to Molina and me. (Rowland 1987; ref (2) is to Molina 1974)

Comparison (co): Cited information provides further explanation by pointing to similarities with other phenomena of the same kind, or by referring to similar results of from a nearby area of research. This latter type of comparison can, of course, be seen as a special, 'weaker' kind of evidence (cf. below). The following passage can be seen as an example of the former kind of comparison:

Molina and Rowland (5) have noted that this stratospheric sink for CFxCl also represents a potential sink for stratospheric O3. This is so because the photodissociation of CFxCl releases chlorine atoms which can catalytically destroy O3 through reactions like those of the nitrogen oxides (NOx) with O3 (5-8) (Cicerone 1974, p. 1165; ref. (5) is to Molina 1974)
Evidence (ev): Information in cited passage is used to prove or support own position. To distinguish it from comparison (above), evidence is used in this case study about citations to information that is directly concerned with the same phenomena as the citing article, that is, information that could be used in an inference to corroborate the main points of the citing article. However, instances of conclusive, logically binding proofs seem to be non-existent in the material. Rather, when information is invoked as evidence, reference is often to the apparent agreement between measurement data and predictions of a theory or a model. The following sentence is a typical example:

    Qualitatively consistent with our predictions are observations of an increase in total O3 in the Northern Hemisphere from 1961 to 1970 (16).

    (Chameides 1975; ref. (16) is to London 1974).

At other times, the evidence invoked through citation is, rather, to a model or process described by the cited article:

    After release of the first chlorine atom by a solar proton, chemical reactions will probably remove the remaining chlorine and fluorine atoms from the CFxCly-1 radical, temporarily forming phosgene-type molecules (5)

    (Cicerone 1974; ref. (5) is to Molina 1974)

Example (ex): Citation points to an instance of a phenomenon intended to illustrate a point in the argument put forward, or is part of an enumeration of such instances, as in the following passage:

    The photochemical processes influencing atmospheric ozone were much more recently elucidated. The original pure-oxygen chemistry proposed by Chapman (1930) was augmented by the recognition of the importance of hydrogen (Hampson, 1964), nitrogen (Crutzen, 1970; Johnston, 1971), and chlorine chemistry (Molina and Rowland, 1974; Stolarski and Cicerone, 1974).

    (Solomon 1988, p 131)

Such an enumeration sometimes, as in the just quoted passage, appears to be a natural part of background information about previous research, typically occurring in the introduction of an article. The example category used here corresponds only roughly to the hierarchical, class-subclass or part-whole relationships in Green (1995b; cf. above section 2.2).

Force (fo): Citation provides explanation by referring to structure or cause of observed phenomena. A typical example in this context is a reference to a chemical reaction, that has been described by the cited work:

    In the stratosphere, ultraviolet radiation dissociates the common fluorocarbons (CFC13 and CF2Cl2), producing free chlorine atoms (1) that destroy ozone by the reaction chain (2)

                Cl + O3  -->  ClO + O2
                ClO + O  -->  Cl + O2

    (Johnston 1980; ref. (2) is to Molina 1974)
Magnitude (ma): Used for citation to measurement data, e.g. in a table. This type of information is often invoked as evidence, but other types of use, such as for employment in a model, also occur. The most obvious example in the material is an article, in which a whole series of measurement data from London (1974) form one row (out of thirteen) in a table of "values of the varios zonally dependent parameters employed in the model" (Temkin 1976, p. 1673).

Method (me): Citation refers to the method employed in the cited work. This does not necessarily mean that same method is employed or even advocated by the citing article. In the instances found in this study rather the opposite is the case, as is seen from the following passage:

The method used in the subjective analysis of London and Kelley (1974) has not been published; because of the judgments involved it would be difficult, if not impossible, to reproduce on a computer. (Heath 1982, p. 7243)

Modification (md): The validity of the cited information is not denied, but qualified further, or its scope is restricted. This can be regarded as a weaker form of objection (cf. below). Yung (1980) may serve as an example, although in this article the modification introduced is not directly associated with the cited document. Molina (1974) is cited in the introduction, as part of the background information. Although no further reference is made to that same source again, the results of the research reported, as described by the abstract, clearly has a bearing on Molina (1974):

The most important finding is that through the reaction BrO + ClO --> Br + Cl + O2 there is a synergistic effect between bromine and chlorine which results in an efficient catalytic destruction of ozone in the lower stratosphere. ... We show from the foregoing that bromine is more efficient than chlorine as a catalyst for destroying ozone... (Yung 1980, p. 339)

Motivation (mo): The function of the citation is to serve as a motivation, in adjunction with other reasons, for the research reported by the citing authors or, more generally, for their writing their paper. This kind of citation should be expected to appear primarily in the introduction of an article, in the statement of purpose of the document, as in this passage:

Ophthalmologists working in equatorial regions have long been familiar with the syndrome of solar conjunctivo-keratopathy. [Description of symptoms follows] ... A new factor has now arisen which threatens significantly to increase this hazard and possibly to extend the geographical area in which this minor but apparently incurable syndrome may be encountered. It is the

---

36 Although the material is too small for any general conclusions, it is interesting to note in this context that the critique sometimes directed against citation analysis for favoring papers on methodology, seems to be completely inappropriate in this case. Only 4 citations, out of a total of 115 (counting multiple references to the same document occurring in one citing article as separate citations), could properly be classified under the category method in this case study. For a more elaborated answer to the accusation against citation analysis of methodology-bias, cf. Garfield (1979, p. 246).
purpose of this note to bring the new circumstances to the attention of ophthalmologists. The new factor is the growing threat to the ozone layer from the ever-increasing quantities of chlorofluoromethane gases released into the atmosphere, mainly from aerosol sprays. Molina and Rowland (1974) described the threat to the ozone layer. (Youngson 1975)

Another example shows how the motivation category also describes research that explicitly sets out to test a hypothesis put forward by a cited article:

**ABSTRACT**

A two-dimensional model is used to predict the 1990 reduction in ozone due to the chlorine compounds formed by chlorofluoromethanes (CFM) photolysis when the CFM release rate is held constant at the 1975 value. ...

1. Introduction

Estimates of the depletion of stratospheric ozone due to chemically active chlorine produced by the photolysis of chlorofluoromethanes (CFM) released into the troposphere have been reported by numerous investigators (e.g., Cicerone et al., 1974; Crutzen, 1974; Molina and Rowland, 1974; Turco and Whitten, 1975; Wofsy et al. (1975). Their calculations suggested that substantial reductions in global ozone (on the order of 3%) could be expected by 1990 if CFM release continues at or at somewhat above the 1974 rates. (Borucki 1980, p. 686)

**Neutralization (ne):** Citation appears to be an attempt at anticipating and disarming possible critique, by means of interpretation of possibly problematic data in line with the authors own argumentation. Consider the following example, the only instance found in the material:

There appears to be general agreement now that the average ozone content of the northern hemisphere rose by 5-7% during the period from 1961 to 1970 (Komhyr et al., 1971; Johnston et al., 1973; London and Kelley, 1974)... One proposed explanation for the increase in the northern hemisphere is that the ozone level there was reduced by extensive atmospheric testing of nuclear weapons during the period 1960-1963 with substantial formation of NO<sub>x</sub> through the heating of air in the nuclear fireballs and that the atmosphere has been recovering from this depletion since the ban on atmospheric testing was established in 1963. (Rowland 1975, p. 32)

**Objection (ob):** Cited information is either overtly refuted or its validity is seriously questioned. As an example of the latter kind of objection, which seems to be the more common, consider the following passage:

Additionally, Rowland and Molina state in their original article (3), "We have not included any estimates for other chlorinated aliphatic hydrocarbons also found in the atmosphere such as CCl<sub>4</sub>... for which there is no evidence for long residence in the atmosphere." Six months later, an article (4) by the same authors was published treating CCl<sub>4</sub> in exactly the same manner that the fluorocarbons were treated in the first article. A month after that,
CH₃Cl, a molecule not even mentioned in the first article, was shown (5) to be the largest single source of chlorine presently established in the atmosphere. Whether or not the authors assumed that fluorocarbons are the only significant source of chlorine available for interaction with ozone in the atmosphere, the fact remains that there are other significant sources than those mentioned in their original article and that they were only pointed out at a later date. There is thus an established record of things having been left out; one may reasonably ask what else has been left out.

(Cairns 1975, p. 1040)

**Quote (qu):** Citation is a direct quotation from cited work. Quotations however, appear to be rare in science, which may not be surprising. In contrast to the humanities, where the research process has been described as a "treasure hunt for mots justes " (Budd 1989, p. 9), science seems to be more concerned with fact gathering and, hence, with the contents of the cited works, rather than with the exact language used. Quotations, when they occur, seldom do so in their own right. They serve some other purpose. In the material studied here only a few instances of direct quotations were found. In two cases, one of which is the passage from Cairns (1975) quoted above, the purpose is obviously to stage an objection. In a third case, which is an instance of self-citation (Rowland 1989), the purpose seems to be that of providing background information, as part of a bibliographical history of the development in the research field. The rest of the quotes are of a chemical reaction formula; the same formula appears also in other articles in the study, but without their giving credit to the likely source from which it has been copied.³⁷

**Result (re):** Used for a relation of implication, viz. "if information Cₒ contained in cited passage is true, and if furthermore conditions C₁, C₂, ... Cₙ hold good, then the consequences will be such and such". Hence, the citing article does not necessarily have to endorse a claim of truth for the cited information; the only claim is for the potential result, given the conditions described by the antecedent of the implication.³⁸ The auxiliary conditions C₁, C₂, ... Cₙ furthermore do not have to be topically related to the cited information. From a strictly logical point of view, the only requirement on Cₒ, C₁, C₂, ... Cₙ taken together is that there must be no contradiction among them. (Since from

³⁷This, admittedly, shows the sometimes arbitrary character of citations; arbitrary in the sense that it does not always seem clear when to add references and when not. It does not mean, however, that citations actually occurring in papers always represent only arbitrary and rather loose links with the cited documents.

³⁸The result (re) category, thus, seems to cover essentially the same type of relationships as the "conditional relevance" described by Cooper (1971):

it is possible for stored information to be relevant to a need only in the presence of other stored information which links it to that need. ... the relevance of a stored statement is conditional on the presence of all other members of its minimal premiss set. One has, therefore, an explication of "conditional relevance" in at least one important sense of that phrase. (op.cit., pp. 25f)
a logical contradiction any proposition can be deduced, in such case the validity of the
results presented by the citing article would be severely put into question.) Several
instances of this type of citation appear in articles from journals that are clearly
peripheral to the field of research concerned with monitoring of the ozone layer. The
SciSearch journal subject categories reveal that they come from such disciplines as
biochemistry and molecular biology, botany, or ophthalmology (cf. the passage quoted
above under motivation from Youngson 1975). It seems natural, that researchers from
"outside" should be more concerned with the implications of the cited information for
their own field of research, rather than with trying to assess the validity of that
information - something that they in most cases probably lack the necessary specialist
competence to be able to do. These citations could be considered, in accordance with
the distinction introduced by Nolin (1995) that was referred to above, as belonging to
the research on the effects of stratospheric ozone depletion, rather than to the field of
research concerned with monitoring. The result category, then, could be a possible
indicator of what Nolin calls "functional integration" (op.cit., p. 244). The following
passage may serve as an example:

Recent studies by Cicerone (4) and Molina and Rowland (7) state that
increased use of fluorocarbons in aerosols and refrigerants could severely
deplete the protective layer of ozone in the stratosphere. This would
increase the level of UV-B radiation reaching the earth's surface. However,
depletion of the ozone layer to the degree postulated (8) would not allow
penetration of irradiation below 290 nm in nature, i.e. irradiation in the UV-
C spectrum. The effect of UV-B (280-320) nm irradiation on plant virus
interactions has not been examined. The object of this study was to
determine the effects of UV-B irradiation on local lesion development of
Chenopodium quinoa Willd. 'Valdivia' plants inoculated with potato virus S
(PVS).

(Semeniuk 1980; ref. (7) is to Molina 1974)

However, instances of the result category are also found in citing articles, topically
more closely related to the cited article, than in the preceeding case. Consider as another
example the following passage:

It has been indicated (1,2) that increasing CF2Cl2 and CFC13 concentrations
to 2 ppb would decrease O3 by 10 percent. A 10 percent decrease in O3, if
distributed uniformly beteween 12 to 40 km, would decrease Ts [=surface
temperature of the Earth] by about 0.07°K

(Ramanathan 1975; ref. (1) is to Molina 1974, (2) is to Cicerone 1974)

No claim is made, as is easily understood from these examples, neither for the mutual
exclusiveness, nor for the joint exhaustiveness of the categories used for classification
of citation functions. This is evident already from the fact that most of the citations in
the case study have been assigned multiple functions. Rather, the categories should be
seen as model examples of relationships holding between citing and cited documents,
explaining at best why a certain document was found useful for or relevant to the
purpose of the citing article. In the next section the results of the analysis undertaken in
the case-study are now finally presented.
5.4 Results

The distribution as for theoretical approach (C="chemical", D="dynamical", U=unspecified) of articles containing citations to London (1974) = L, to Molina (1974) = M, and to both (L&M) is shown in Table 2. (That is, articles containing citations to both London (1974) and Molina (1974) are not contained in the figures of the second and third columns.):

<table>
<thead>
<tr>
<th>Theory</th>
<th>L</th>
<th>M</th>
<th>L&amp;M</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-</td>
<td>27</td>
<td>2</td>
</tr>
<tr>
<td>D</td>
<td>8</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>U</td>
<td>8</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>Σ</td>
<td>16</td>
<td>45</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 2: Distribution of citing articles over theoretical approach

The distribution over functions of the citations to London (1974) = L and Molina (1974) = M is now shown in Table 3 on the next page. Figures represent number of citations, if nothing else is stated. It should be kept in mind that some articles contain several citations to the same source and that the total number of citations to either of the two articles (bottom row) is less than the sum of all classified instances, since many citations have received multiple classifications. The first column contains the various citation functions. The columns headed by L tot. and M tot. give the total number of citations to London (1974) and Molina (1974) respectively, that have been assigned the function in question. LC, LD, LU and MC, MD, MU represent articles of "chemical" (C), "dynamical" (D) and unspecified (U) theoretical approach respectively that cite either London (1974) or Molina (1974). (That is, for each citation function, L tot. = LC + LD + LU, and M tot. = MC + MD + MU). L med. and M med. finally represent the median publishing years for documents containing citations to London (1974) and Molina (1974) respectively, that serve the specified function.
Table 3: Distribution of citation functions:

<table>
<thead>
<tr>
<th>F</th>
<th>L tot.</th>
<th>L C</th>
<th>L D</th>
<th>L U</th>
<th>L med</th>
<th>M tot.</th>
<th>M C</th>
<th>M D</th>
<th>M U</th>
<th>M med</th>
</tr>
</thead>
<tbody>
<tr>
<td>ba</td>
<td>10</td>
<td>-</td>
<td>2</td>
<td>8</td>
<td>1979.5</td>
<td>47</td>
<td>26</td>
<td>5</td>
<td>16</td>
<td>1980</td>
</tr>
<tr>
<td>bi</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>6*</td>
<td>6*</td>
<td>-</td>
<td>-</td>
<td>1987.5</td>
</tr>
<tr>
<td>co</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>1983</td>
<td>2</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>1984.5</td>
</tr>
<tr>
<td>ev</td>
<td>10</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>1979</td>
<td>16</td>
<td>14</td>
<td>-</td>
<td>2</td>
<td>1975</td>
</tr>
<tr>
<td>ex</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1980</td>
</tr>
<tr>
<td>fo</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>13</td>
<td>11</td>
<td>-</td>
<td>2</td>
<td>1984</td>
</tr>
<tr>
<td>ma</td>
<td>10</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>1980.5</td>
<td>5</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>1982</td>
</tr>
<tr>
<td>me</td>
<td>4</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>1982.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>md</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1974</td>
</tr>
<tr>
<td>mo</td>
<td>3</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>1980</td>
<td>13</td>
<td>8</td>
<td>1</td>
<td>4</td>
<td>1980</td>
</tr>
<tr>
<td>ne</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1975</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ob</td>
<td>5</td>
<td>-</td>
<td>1</td>
<td>4</td>
<td>1982</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1975</td>
</tr>
<tr>
<td>qu</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>1977.5</td>
</tr>
<tr>
<td>re</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1975</td>
<td>10</td>
<td>8</td>
<td>-</td>
<td>2</td>
<td>1977.5</td>
</tr>
<tr>
<td>Total</td>
<td>29</td>
<td>4</td>
<td>10</td>
<td>15</td>
<td>1980</td>
<td>86</td>
<td>60</td>
<td>8</td>
<td>18</td>
<td>1980</td>
</tr>
</tbody>
</table>

It should be observed that the figures in Table 3 for citations to Molina (1974) representing the functions bibliography (bi) have been marked by asterisks to indicate that they are all instances of self-citations. The figures for background (ba), evidence (ev) and magnitude (ma) are also inflated to a certain extent by the presence of self-citations to Molina (1974) in the material, as can be seen from Table 4 (below). This fact most likely explains the apparent "theory-dependence" (if by "theory-dependence"...
is meant a high degree of correspondence between theoretical approach of citing and cited articles) of such functions as bibliography (bi) and magnitude (ma). It is at least difficult to find any other plausible reason why these categories, both of which represent contingent, neither logical nor causal connections, should show an even stronger "theory-dependence" than such categories as evidence (ev), force (fo) or result (re).40

Apart from these "anomalies", however, it is interesting to note further in Table 3 that evidence (ev) seems to be more "theory-dependent" than for example background (ba), something that appears to accord well with intuition, given the explications of these two citation functions in the preceeding section. The two instances of London (1974) serving as evidence (ev) in a C-article, containing also citations to Molina (1974) that likewise serve as evidence, are both from an extensive review-article (in Spanish). The particular points in London (1974) referred to in this article, about the decrease in total ozone between 1957-61 being followed by an increase during the period of 1962 -1970 and the marked difference in this respect between the Northern and the Southern Hemispheres, are further interpreted as indicating a return of the ozone to its normal value after the earlier intensive nuclear bomb tests in 1952-1962, an interpretation that is by no means inherent in London (1974). Apart from these two "exceptions" in the case of London (1974), we observe in the case of Molina (1974) that citations serving as evidence do not occur in articles of an opposite theoretical approach than that of the cited source.

<table>
<thead>
<tr>
<th>Function</th>
<th>M self</th>
</tr>
</thead>
<tbody>
<tr>
<td>ba</td>
<td>9</td>
</tr>
<tr>
<td>bi</td>
<td>6</td>
</tr>
<tr>
<td>co</td>
<td>-</td>
</tr>
<tr>
<td>ev</td>
<td>7</td>
</tr>
<tr>
<td>ex</td>
<td>-</td>
</tr>
<tr>
<td>fo</td>
<td>1</td>
</tr>
<tr>
<td>ma</td>
<td>3</td>
</tr>
<tr>
<td>me</td>
<td>-</td>
</tr>
<tr>
<td>md</td>
<td>-</td>
</tr>
<tr>
<td>mo</td>
<td>1</td>
</tr>
<tr>
<td>ne</td>
<td>-</td>
</tr>
<tr>
<td>ob</td>
<td>-</td>
</tr>
<tr>
<td>qu</td>
<td>1</td>
</tr>
<tr>
<td>re</td>
<td>-</td>
</tr>
<tr>
<td>Tot. citations</td>
<td>22</td>
</tr>
</tbody>
</table>

Table 4: Citation functions for self-citations to Molina (1974)

40 The self-citations do not, however, affect the median publishing years for either background or evidence.
The seven articles in the sample containing self-citations to Molina (1974), of which F.S. Rowland is co-author, were also treated separately as a special case. The distribution over citation functions for them is shown in Table 4 (above). Apart from the bibliography (bi) function appearing as a unique category among the self-citations, the relatively high frequency of those serving as evidence (ev) is noteworthy, as compared with the rate of "evidential" citations to Molina (1974) in the material as a whole. One may also observe that the average number of self-citations per citing article appears to be higher than normal; approximately 3.14, compared to 1.42 for the citations to Molina (1974) in the sample that are not self-citations, and 1.26 for the citations to London (1974).

Tables 5a-5b below show the number of shared descriptors and union overlap measures or similarity scores, sim (X,Y), for those articles of the original sample that cite London (1974) and Molina (1974) respectively, for which records were retrieved in Aerospace database (Dialog file 108). N is the number of references to the cited article, London (1974) or Molina (1974), T is again the theoretical approach, Functions are the citation function categories assigned (in cases where several citations of a document have been assigned the same category, the number of citations of each category is given within parentheses. Where only one category is used, the number of citations of that category naturally equals N). The number of shared descriptors in Aerospace database of citing source with London (1974) and Molina (1974), is shown in the columns headed by nL and nM respectively. That is, to allow for comparison of topical relatedness with both cited and non-cited document, the number of shared descriptors and similarity scores for both London (1974) and Molina (1974), sim (X,L) and sim (X,M) respectively, are given for each source, irrespective of which of these two it is citing. The columns headed by uL and uM finally give the denominators in the union overlap measure formula, sim (X,Y) = n (X ∩ Y) / n (X ∪ Y), for Y = L, London (1974), and Y = M Molina (1974) respectively.

Table 5a (next page) shows that for London (1974) a topic matching relationship, i.e. two or more shared descriptors, is holding between citing and cited source in 8 out of 13 instances. On the other hand, such a relationship is also holding in 6 instances with Molina (1974), which is not cited by these sources. Again, in 8 out of 13 instances - more than half of the cases, the topical relatedness of the sources in Table 5a (or the absence of such a relatedness, as in the case of Temkin 1975, 1976) with the non-cited document, Molina (1974), appears to be on the same level or even higher, than that of the cited one, London (1974). In other words, the fact that a document cites another document does not appear to make very much difference when it comes to determining the topical relatedness of the two, at least not as measured by the criterion for topic matching used here or by the union overlap measure.
Table 5a: Sources citing London (1974), but not Molina (1974), in Aerospace db
(Py) = publishing year, N = number of references in source to cited article, T= theoretical approach, \( \cap L, \cap M \) = numerators and \( \cup L, \cup M \) = denominators in the union overlap measure formula, \( \text{sim} (X,Y) = n(X \cap Y) / n(X \cup Y) \), for London (1974) and Molina (1974) respectively.

<table>
<thead>
<tr>
<th>Source X (py)</th>
<th>N</th>
<th>T</th>
<th>Functions</th>
<th>( \cap L )</th>
<th>( \cup L )</th>
<th>\text{sim} (X,L)</th>
<th>( \cap M )</th>
<th>( \cup M )</th>
<th>\text{sim} (X,M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crutzen (75)</td>
<td>1</td>
<td>U</td>
<td>ba</td>
<td>3</td>
<td>19</td>
<td>0.158</td>
<td>2</td>
<td>20</td>
<td>0.100</td>
</tr>
<tr>
<td>Temkin (75)</td>
<td>1</td>
<td>D</td>
<td>ma</td>
<td>0</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>Temkin (76)</td>
<td>1</td>
<td>D</td>
<td>ma</td>
<td>0</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td>Bauer (79)</td>
<td>1</td>
<td>U</td>
<td>ba</td>
<td>2</td>
<td>18</td>
<td>0.111</td>
<td>3</td>
<td>17</td>
<td>0.176</td>
</tr>
<tr>
<td>Callis (79)</td>
<td>1</td>
<td>D</td>
<td>ev, ma</td>
<td>2</td>
<td>19</td>
<td>0.105</td>
<td>3</td>
<td>18</td>
<td>0.167</td>
</tr>
<tr>
<td>Chang (79)</td>
<td>3</td>
<td>U</td>
<td>ba</td>
<td>2</td>
<td>18</td>
<td>0.111</td>
<td>2</td>
<td>18</td>
<td>0.111</td>
</tr>
<tr>
<td>Vernazza (80)</td>
<td>1</td>
<td>U</td>
<td>ba, mo</td>
<td>2</td>
<td>14</td>
<td>0.143</td>
<td>0</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>Natarajan (81)</td>
<td>2</td>
<td>D</td>
<td>ba, ev, ob</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>2</td>
<td>18</td>
<td>0.111</td>
</tr>
<tr>
<td>Heath (82)</td>
<td>1</td>
<td>U</td>
<td>me, ob</td>
<td>4</td>
<td>17</td>
<td>0.235</td>
<td>1</td>
<td>20</td>
<td>0.050</td>
</tr>
<tr>
<td>Angell (83)</td>
<td>1</td>
<td>D</td>
<td>ev, ma</td>
<td>2</td>
<td>17</td>
<td>0.118</td>
<td>2</td>
<td>17</td>
<td>0.118</td>
</tr>
<tr>
<td>Mastenbrook (83)</td>
<td>1</td>
<td>D</td>
<td>co, ev, ma</td>
<td>0</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>Wilcox (83)</td>
<td>1</td>
<td>U</td>
<td>me, ob</td>
<td>3</td>
<td>16</td>
<td>0.188</td>
<td>0</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>Towe (92)</td>
<td>1</td>
<td>D</td>
<td>ev, ma</td>
<td>1</td>
<td>14</td>
<td>0.071</td>
<td>0</td>
<td>15</td>
<td>0</td>
</tr>
</tbody>
</table>

The correlation between citation and topical relatedness appears to be higher for sources that cite Molina (1974), but not London (1974), judging from the figures in Table 5b (next page). But still, in more than half of the instances the requirement for topic matching is fulfilled also for the non-cited London (1974). And in both Tables 5a and 5b there are some obvious "anomalies", that is documents that show lower similarity scores with the cited article than with the non-cited one.
<table>
<thead>
<tr>
<th>Source X (py)</th>
<th>N</th>
<th>T</th>
<th>Functions</th>
<th>( \cap L )</th>
<th>( \cup L )</th>
<th>sim ( (X, L) )</th>
<th>( \cap M )</th>
<th>( \cup M )</th>
<th>sim ( (X, M) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cicerone (74)</td>
<td>6</td>
<td>C</td>
<td>co, ev(2), fo(4), qu, re</td>
<td>1</td>
<td>19</td>
<td>0.053</td>
<td>4</td>
<td>16</td>
<td>0.250</td>
</tr>
<tr>
<td>Lovelock (74)</td>
<td>2</td>
<td>D</td>
<td>md, ob</td>
<td>2</td>
<td>19</td>
<td>0.105</td>
<td>3</td>
<td>18</td>
<td>0.167</td>
</tr>
<tr>
<td>Lubkin (75)</td>
<td>1</td>
<td>C</td>
<td>ba, re</td>
<td>2</td>
<td>18</td>
<td>0.111</td>
<td>4</td>
<td>16</td>
<td>0.250</td>
</tr>
<tr>
<td>Ramanathan (75)</td>
<td>5</td>
<td>C</td>
<td>ba(2), re(3)</td>
<td>1</td>
<td>22</td>
<td>0.045</td>
<td>2</td>
<td>21</td>
<td>0.095</td>
</tr>
<tr>
<td>Rowland (76a)</td>
<td>4</td>
<td>C</td>
<td>ba, ev(2), ma, mo</td>
<td>3</td>
<td>19</td>
<td>0.158</td>
<td>4</td>
<td>18</td>
<td>0.222</td>
</tr>
<tr>
<td>Rowland (76b)</td>
<td>1</td>
<td>C</td>
<td>ba</td>
<td>2</td>
<td>19</td>
<td>0.105</td>
<td>4</td>
<td>17</td>
<td>0.235</td>
</tr>
<tr>
<td>Crescentini (79)</td>
<td>1</td>
<td>C</td>
<td>ba, mo</td>
<td>1</td>
<td>15</td>
<td>0.067</td>
<td>2</td>
<td>14</td>
<td>0.143</td>
</tr>
<tr>
<td>Borucki (80)</td>
<td>1</td>
<td>C</td>
<td>ba, mo</td>
<td>2</td>
<td>19</td>
<td>0.105</td>
<td>3</td>
<td>18</td>
<td>0.167</td>
</tr>
<tr>
<td>Chameides (80)</td>
<td>1</td>
<td>C</td>
<td>ba, mo</td>
<td>2</td>
<td>22</td>
<td>0.091</td>
<td>3</td>
<td>21</td>
<td>0.143</td>
</tr>
<tr>
<td>Goldan (80)</td>
<td>1</td>
<td>C</td>
<td>ba</td>
<td>1</td>
<td>20</td>
<td>0.050</td>
<td>1</td>
<td>20</td>
<td>0.050</td>
</tr>
<tr>
<td>Johnston (80)</td>
<td>1</td>
<td>U</td>
<td>ba, fo, qu</td>
<td>1</td>
<td>19</td>
<td>0.053</td>
<td>3</td>
<td>17</td>
<td>0.176</td>
</tr>
<tr>
<td>Shapiro (80)</td>
<td>1</td>
<td>D</td>
<td>ba, ex, mo</td>
<td>2</td>
<td>19</td>
<td>0.105</td>
<td>1</td>
<td>20</td>
<td>0.050</td>
</tr>
<tr>
<td>Yung (80)</td>
<td>1</td>
<td>U</td>
<td>ba, md</td>
<td>3</td>
<td>17</td>
<td>0.176</td>
<td>3</td>
<td>17</td>
<td>0.176</td>
</tr>
<tr>
<td>Trainer (83)</td>
<td>1</td>
<td>U</td>
<td>ba, mo</td>
<td>1</td>
<td>20</td>
<td>0.050</td>
<td>2</td>
<td>19</td>
<td>0.105</td>
</tr>
<tr>
<td>Cicerone (87)</td>
<td>1</td>
<td>C</td>
<td>ba, ex</td>
<td>2</td>
<td>17</td>
<td>0.118</td>
<td>4</td>
<td>15</td>
<td>0.267</td>
</tr>
<tr>
<td>Barrett (88)</td>
<td>1</td>
<td>C</td>
<td>ba, fo</td>
<td>2</td>
<td>16</td>
<td>0.125</td>
<td>1</td>
<td>17</td>
<td>0.059</td>
</tr>
<tr>
<td>Solomon (88)</td>
<td>1</td>
<td>U</td>
<td>ba, ex</td>
<td>1</td>
<td>20</td>
<td>0.050</td>
<td>0</td>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>Rowland (89)</td>
<td>2</td>
<td>C</td>
<td>ba, bi, fo, ma, qu</td>
<td>1</td>
<td>21</td>
<td>0.048</td>
<td>3</td>
<td>19</td>
<td>0.158</td>
</tr>
<tr>
<td>Stolarski (92)</td>
<td>1</td>
<td>C</td>
<td>ba</td>
<td>1</td>
<td>15</td>
<td>0.067</td>
<td>1</td>
<td>15</td>
<td>0.067</td>
</tr>
</tbody>
</table>

**Table 5b:** Sources citing Molina (1974), but not London (1974) in Aerospace db. 

(Py) = publishing year, \( N \) = number of references in source to cited article, \( T \) = theoretical approach, \( \cap L \), \( \cap M \) = numerators and \( \cup L \), \( \cup M \) = denominators in the union overlap measure formula, \( \text{sim}\ (X, Y) = \frac{n(X \cap Y)}{n(X \cup Y)} \), for London (1974) and Molina (1974) respectively.

Only three records for articles citing both London (1974) and Molina (1974) were found in Aerospace database; a number evidently too small to permit any interesting conclusions.
6 Discussion

The high rate of articles that had to be left unspecified (U) with respect to theoretical approach makes it difficult to interpret the results manifested through Tables 2 and 3 in section 5.4. This is particularly true for the figures representing citations to London (1974). It may perhaps be seen as an indication that this article is itself more "neutral" in its theoretical approach, than Molina (1974). The contents of London (1974) are such that they do not require the endorsement of any particular theory; essentially the article is a report on measurement data. Thus, it is possible that London (1974) should itself have been left unspecified, rather than being classified as "dynamical". Interpreting the results in this way, there would still seem to be a slight tendency in the sample towards "theory-dependency", in the sense that an article is more likely to receive citations from another article of the same theoretical approach as itself, than from an article of an opposite theoretical approach.

Such an interpretation would appear to gain support from the study by Cano (1989, p. 286), which showed a low "utility level", as judged by the citing authors themselves, for "negational" references, i.e. references the correctness of which is disputed (cf. Moravcsik 1975, p. 88), and which, consequently, is more likely to have a different theoretical approach than that of the citing paper. "Negational" references, in the same sense, were also found to be considerably less frequent than "confirmative" or "affirmative" in the studies of Moravcsik (op.cit.) and Chubin (1975) - the latter of finding as much as 95% of references studied to be "affirmative" (op.cit., p. 428), although no difference in this respect was found by Cano (1989). The results of the present study also reveal, as shown in Table 3, that only 14 of a total of 115 citations (approximately 12%) belong to the three categories used for classification of "negational" references, viz. modification (md), neutralization (ne) and objection (ob). Thus, also in this respect the results of the present study would seem to agree with those of Moravcsik (op.cit.) and Chubin (op.cit.).

In a study of references of articles in high energy physics, Moravcsik (1975) observed that about 40% of the references appeared to be "perfunctory", rather than "organic". By a "perfunctory" reference was meant one that is serving mainly "as an acknowledgment that some other work in the same general area ha been performed" (op.cit., p. 88). It may be difficult to compare, but references judged "perfunctory" would seem to cover roughly the same kind of citations as those that were classified as providing background information (ba) in the present study. Thus, in this respect, given the relatively high proportion of instances of this class, 34% of the citations to London (1974) and 55% of those to Molina (1974), the results of this study would seem to confirm those of Moravcsik (1975). However, the pre-dominance in the present case-study of citations classified as providing background information should not be attributed undue significance. First, it must be remembered that multiple classification was allowed and that some of the citations classified as (ba) simultaneously were seen to serve other functions as well. Secondly, and more important, a citation to background information does not necessarily indicate information of less relevance, not even in a logical,
evidential sense of the notion of relevance. The philosopher of science Dudley Shapere (in Suppe 1977; cf. Philipson 1984) has pointed to the important role of background information in science as a justification for abstraction and modeling, giving as a classic example Bohr's theory of the structure of an atom, which in its original version disregarded both the movements of the atomic nucleus and possible effects of the high velocity of the electrons surrounding it. The justification for this was given by the background information provided by the classical theory of electricity and by the special relativity theory. There may be reason to suspect, that this kind of background information is often never even cited in the scientific journals. The readership of scientific journals could, in general, be assumed to have a certain familiarity with the questions treated in the articles, particularly when we talk about the readership of journals specializing in one single field of research. This ought to mean, that a certain level of background knowledge is taken for granted, both by the editors and by the contributors of papers to the journal. A probable effect thereof is that a good number of potential references are judged unnecessary for the purpose of stating the argument of a paper. Possibly this is one of the factors responsible for the phenomenon of "obliteration" (cf. section 3.3 above).

In conjunction with this one might also speculate upon another possible explanation for the apparent "theory-dependence" of citations in the case-study. Should it prove true that "chemists" are in fact less prone to cite "dynamists", than they are to cite other "chemists", perhaps one of the reasons could be that they possess a common "background knowledge", according to which the dynamical effects of the circulation of air are legitimately negligible in the the theoretical models of stratospheric ozone distribution, and that taking these effects into account would only complicate things unduly. This "background knowledge", however, seldom appears in the form of overt citations, although it is "interpretatively relevant" in the negative sense, that it tells the chemically inclined scientists what is safely ignorable in the interpretation of data. It is part of "the background of unquestionable and unquestioned familiarity which is simply taken for granted"(Schutz 1970, p. 27; cf. above section 2.2).

In another sense it is difficult to interpret the figures representing the distribution of citation functions, shown in Table 3. For example, evidence (ev) and method (me) are functions that intuitively would seem to account for a high relevance of the cited document to the citing one, and that would presumably serve as likely explanations of a high rate of citedness. That is, it was expected that these important functions would be more frequently served by citations to Molina (1974), thus providing one possible explanation for its substantially higher rate of citedness as compared with London (1974). Yet, judging from the results of our analysis, both evidence (ev) and even more so method (me) appear to be more frequent as functions served by the citations to London (1974). On the other hand, functions that intuitively would also seem to carry a heavy weight of relevance, such as motivation (mo) and result (re), do appear to be more frequent among citations to Molina (1974), as expected.

Difficult to explain is also the result that the median publishing year of articles invoking Molina (1974) as evidence (ev) is 1975, whereas for those that cite London (1974) it is
1979. There would seem to be good reason to expect a longer life for the evidential relevance of a highly cited article, representing a victorious theory, than for a "looser" with a much lower degree of citedness. This result could of course be the effect of a poor selection of articles that cite Molina (1974). With a larger, more representative selection the median publishing year for "evidential" citations to Molina (1974) would perhaps approach that of London (1974). Still, bearing in mind that the overall median publishing year in the present sample of articles is the same for those that cite Molina (1974), as for those that cite London (1974), i.e. 1980, it is not necessarily true that a larger and better selection of articles would automatically prolong the "evidential relevance" of Molina (1974).

Other results concerning the "time-dependences" of citation functions, as manifested through the median publishing-years in Table 3, are perhaps easier to explain. For example, it seems reasonable to expect that a document may be cited for background information (ba) during a longer period of time, than it will continue to be invoked as evidence (ev). As science develops, theories change and new facts are discovered, what will be counted as valid evidence will change too. But even long since refuted hypotheses may still be considered relevant to the purpose of tracing the historical development within a field of research or just "setting the scene" before presenting the results of ones own research. It may also be that results of earlier research, although still considered as valid, will not be invoked so often as evidence simply because of their age, and that there is a general tendency in science, when stating ones case, to seek support rather from the newest possible sources. This could in fact be a possible explanation for the "anomaly" noted above concerning the apparently more shortlived "evidential relevance" of Molina (1974). Since the authors of that article continued to develop constantly their theory, publishing a multitude of articles during a succession of years, it may well be that other scientists in the field of research, when making reference to the CFC-theory, preferred to cite instead these other articles, as the most recent evidence available.

The fact that citations serving the function of explanatory force (fo) seem to have a relative "longer life" than most other functions is also not surprising. The reference to a chemical reaction or structure should be of such permanence that it could be expected to be found not only in articles in scientific journals, but even in textbooks. Essentially the same would seem to be valid also for citations serving a bibliographical (bi) function, but since the few instaces of this function in the material are all self-citations, it is probably wise to abstain from drawing any conclusions whatsoever about them. The methodological (me) citations in the material are also few in number, but nevertheless their relatively "long life" could prove to be more than just an accidental effect of the selection. If so, support could be gained from the results of a study by Small (1977), showing, by means of co-citation clustering, how a scientific paper that was formerly frequently cited for "theoretical" reasons as describing the structure of collagen,41

---

41 Garfield (1979) describes collagen as "a large, triple-helical, protein molecule in the human body whose functional role is to form fibers that attach to tendons and joints and
suddenly ceased to be among the highly cited papers for a short time, when the focus of research in the field shifted from structural studies to biosynthesis, only to reappear as one of the high ranking cited sources a year later, but then cited rather for its methodology (cf. Garfield 1979, p. 127f.).

It does not seem possible to detect any general correlation between citation function and topic matching or similarity score in Tables 5a-5b. On the one hand, there are a few instances for which the level of topical relatedness seems to accord well with citation functions and with reason, for example the two citations by Temkin (1975, 1976) to London (1974) in Table 5a, classified as magnitude (ma). The sole function of these citations is to provide measurement data. Evidently, there is no necessary topical relatedness between two documents connected in this very loose way. (Just think of all those documents that cite statistical figures from sources such as almanacs and yearbooks.) In these instances, thus, a low topical relatedness would seem justifiable. But, on the other hand, take the citations the sole function of which is to serve as background information (ba), for example. There are six sources containing such citations in Tables 5a and 5b, three of them citing London (1974), the other three citing Molina (1974). Four of these fulfil the requirement stated for topic matching, two do not. But again, the act of citation does not by itself appear to make any substantial difference. In the same cases, when the requirement for topic matching is fulfilled, it is fulfilled for cited and non-cited article alike.

We have already in chapter 4 touched briefly upon one possible reason for the relatively weak correlation between citations and topic matching. A citation refers most often only to a particular point of the cited document. Descriptors also, taken individually, often describe only part of the contents of a document. Taken together they should ideally describe the whole document, but that is rarely the case. The art of indexing is, rather, an art of selection, drawing out the essentials of a document and classifying these essential points in accordance with a thesaurus. However, it is not difficult to imagine a citation that is not to one of these essential points, but rather to information that may appear to be peripheral in the context of the original document, but that is indeed essential in the context of the citing article. Or, turning things the other way around, it is possible to think of a citation that is indeed to a central point in the cited document, but which is not part of one of the central points of the citing document. A possible instance of the latter type is Mastenbrook (1983). As can be seen in Table 5a, this article has a similarity score of 0 with London (1974) - not only in Aerospace database, but also in the Pascal database, using exclusively English descriptors for the computation. However, its citation link with London (1974) appears to be rather strong, with the citation providing both measurement data (ma), functioning as an item of comparison (co) and lending supporting evidence (ev) to the conclusion that "the long-term trend in stratospheric water and its similarity to the long-term trend in stratospheric ozone suggest that these changes arise from long-term changes in the intensity of the circulation" (Mastenbrook 1983, p. 2164). But obviously, the article is not about line arterial walls. The primary constituent of connective tissue, it is the most abundant protein in the body." (op. cit., p. 123)

There are several "anomalies" in Tables 5a and 5b other than that exemplified by the case of Mastenbrook (1983). By "anomalies" is meant, again, those instances in which the non-cited article proved to have a higher similarity score with the non-cited article, than with the cited one; i.e., in Table 5a, Bauer (1979), Callis (1979), Natarajan (1981), Mastenbrook (1983), further, in Table 5b, Shapiro (1980), Barrett (1988) and Solomon (1988). The first three of these proved to be the only three instances in the sample of citing articles indexed in Aerospace database by the descriptor "ultraviolet radiation", which they share with Molina (1974), but not with London (1974). Callis (1979) and Natarajan (1981) are also closely linked, since they are both co-authors of each other. For the three "anomalous" articles that cite Molina (1974) it is more difficult to find any similarly discriminating factor. True, Shapiro (1980), Barrett (1988) and Solomon (1988) all share the descriptor "atmospheric composition" with London (1974), which is not shared by Molina (1974), but that is the case also with another 7 of the articles that cite Molina (1974), which are not "anomalous" in this sense. Neither do we seem to get any indication of some futher common property of these three articles that could possibly explain their differing status by comparing the theoretical approach or citation functions assigned to them. It seems we have to content ourselves with the observation that in all three instances the citation link with the cited source appears to be rather "weak", serving principally the function of providing background information (ba), occurring typically in the introductory part of the articles, with the auxiliary functions of giving examples (ex), motivating purpose (mo) or providing explanatory force (fo).

The complex relationship between citation and topical relatedness, however, is reflected not only through these apparent "anomalies", but also in cases where, prima facie, there seems to be a strong correspondence between the two. To take just one example, consider the case of Heath (1982) in Table 5a. This article might be called a "methodology" paper. It is entitled "On the adequacy of the fixed locations of the surface-based international network for inferring interannual ozone variability" and is concerned with the methods of measurement and analysis of ozone data. In this it differs somewhat in its focus from London (1974), which is principally concerned only with analysis, taking the methods of measurement more or less for granted in its relying on the surface-based international network of local stations, that use either so-called Dobson spectrophotometers or optical filter techniques to measure ozone. Heath (1982) contains only one citation to London (1974), quoted above (section 5.3) as an example of the use of the methodology (me) category. Apparently, Heath (1982) is fairly critical towards the method used by London (1974). This is reflected indirectly also in the conclusions of the article, which point to the inaccuracies inherent in measurements from the ground-based network, resulting from difficulties in "properly weighting changes in different latitude zones, and inadequate network coverage in some areas of the globe" (Heath 1982, p. 7251). To overcome these problems, the recommendation is to compare surface-based measurements with results "based on more nearly complete
measurements, such as satellite data" (ibid.). Clearly, although the citation to London (1974), as seen in isolation, does not necessarily seem to indicate a close relationship between the two articles, the conclusions of Heath (1982) have a definite bearing on the information contained in London (1974) and would presumably be of help to a potential reader of the latter article in evaluating the information contained therein. Is this also the reason for the relatively high similarity score for this pair of documents? If we examine the case we find that the descriptors that both documents have in common in Aerospace database (Dialog file 108) are four: atmospheric composition, ozone, Northern Hemisphere, Southern Hemisphere. In the Pascal database (Dialog file 144) they have three descriptors in common: total ozone content, Northern Hemisphere, Southern Hemisphere. The fact that the similarity score between London (1974) and Heath (1982) is actually higher - 0.250, in Pascal database, than it is in Aerospace database - 0.235, is a result of fewer descriptors being assigned to London (1974) in the former database, which makes the denominator of the quotient in the union overlap measure smaller.42 In summary, in the case of Heath (1982), the citation link in itself seems to be rather "weak", while the categories used to classify this linkage do appear to capture the "relevance relationship" between the two documents. On the other hand, while the topical relatedness with London (1974), as reflected by the number of shared descriptors and by the union overlap measure, in comparison with the corresponding measures for other documents obviously is rather "strong", the index terms on which this "strength" is based do not seem to reflect accurately the importance or relevance of Heath (1982) to London (1974).

What are we to make of the topic matching relationships investigated, then? Whatever the explanation, it seems clear from the examples given that at least in some cases they do not give an adequate picture of the relevance of the information contained in one document to the purpose of another. The weak correlation observed in this study between citations and topical relatedness seems to accord well with the results of Harter (1993), who found a mean similarity score (union overlap measure) between citing and cited library journal articles ranging somewhere between 0.15-0.20. The same study also comprised a "macroanalysis" of semantic relationships between citing and cited documents, involving comparison of LC class numbers. A conclusion was that both the "microanalysis", i.e. the similarity scores, and the "macroanalysis" provided "some evidence that the subject similarity among pairs of cited and citing documents is frequently small. Often the citing and cited documents are in different fields altogether" (op. cit., p. 550).

42 This points to another weakness of this measure: given a case in which two documents have at least one index term in common, the fewer index terms each document has been assigned, the higher the similarity score will be. It is possible to imagine a case in which only one very general descriptor is assigned to each document; if that general descriptor happens to be the same for both documents, the similarity score will be 1.0 - the highest possible. This is in fact the case, as noted earlier, with London (1974) and Molina (1974), as indexed in EMBASE, where both documents were assigned the one descriptor environmental health. Cf. above note 24.
However, both Harter (1993) and Green (1995a, p. 649) are careful to point out the weaknesses of the union overlap similarity score as a measure of topical relatedness, e.g. its vulnerability to interindexer inconsistency. This and other possible shortcomings of this measure that have been noted above affect, of course, also our criterion for topic matching. To be sure, there are other possible measures of topical relatedness, some of which have been briefly touched upon. It may well be that some of those methods, or a combination of them would prove better suited to describe the nature of the relevance relationships involved. However, the method chosen appears to reflect a "real life" subject search better, than for example term co-occurrence, since it is directed exclusively to the most typically topic-oriented field of a database, the descriptor-field. The fact that this field contains controlled vocabulary, gave reason to expect that it would be better suited to express relationships between documents, than, e.g., titles or abstracts. This assumption, of course, may prove to be mistaken. The results of the present study, however, does not give enough evidence for rejecting it out of hand.

Finally, how should the relationship between the citation functions studied and the concept of relevance be assessed? We have already in section 2.2 anticipated this discussion, for example by the observation that citations serving the function of evidence, can sometimes be seen as corresponding to the type of structural relevance described by Green (1995b) in terms of components or "slots in a framework" (op. cit., p. 660). The results of the case-study indicate, further, that evidence is indeed an important function, served by roughly 34% of the citations to London (1974) and by nearly 20% of those to Molina (1974), as can be seen from Table 3. However, it is by no means the most frequent and obviously not the only important function served by citations. But according to the "logical view" of relevance expounded in chapter 2, the evidential function is part of the very definition of relevance. Does this mean that we should reject citation analysis and the classification of citation functions as a viable method for understanding what relevance is about? Not necessarily. It all depends on whether we are prepared to accept ready-made definitions of the concept as valid per se, judging our empirical results with these definitions as measure-sticks, or whether we are willing to take an unconditional approach, trying to reconstruct the concept of relevance instead from the results of empirical investigations. Although to this author the latter alternative clearly seems preferable, it must be admitted that such a reconstruction would take a more comprehensive study than this, encompassing a larger material from several different fields of research, involving also the social sciences and the humanities.

The apparent "theory-dependence" of citations may in fact seem natural, if we agree with Foskett (1972) that "relevance means being part of the paradigm, or public knowledge, or consensus in a field" (op. cit., p. 78). The relevance of citations would partly consist, then, in providing a scientist working within a certain "paradigm" with the necessary theoretical equipment for the solving of "puzzles" within that same paradigm. At the same time, the citations will also serve to convince the reader of a scientific paper, that the research reported therein does indeed conform to the generally accepted requirements that make up the paradigm; that is, the citations will serve as evidence, in a more general sense, of the fact that the research reported builds upon
what is already known and accepted. By the same token, the distinction introduced between citation functions and citation motivations could possibly also be seen as corresponding to the earlier distinction between relevance and pertinence, in the interpretation proposed by Foskett (ibid.) as expressing a difference between public and private knowledge (cf. sections 2.2 and 3.2 above). The suggestion would be, then, that a study of citation functions is perhaps better suited in order to acquire knowledge about relevance relationships, whereas surveys or interviews made in order to get at possible motivations for citations are more likely to produce insights about the reasons why information is considered pertinent to an individual scientist's needs. To develop this suggestion into a substantial claim, however, would take an extensive comparative study, employing both these methods on the same material, and involving also the social sciences and the humanities.

Although today, as noted earlier (cf. section 2.1), no single view of relevance can lay claim to general validity, the method of uncovering relevance relationships by means of a description of citation functions would appear to fit best with a "pragmatist view" of relevance, which is concerned primarily with the utility of information. Excepting those hopefully rare cases in which an author is just copying the references of another source without consulting the primary source, citations can be seen as instances of "recorded use". Now, it has been objected that "a potential instrumental relationship cannot be extracted from a prior instrumental relation" (Hjarland 1992, p. 180). This may seem to indicate, that even if the citation function categories used here would be recognized as describing also some kind of relevance relationships, citation analysis could not be used for projecting future relevance relationships. However, this objection concerns primarily only the "subject-relatedness" (ibid.) of documents that are linked by citation. And as we have seen, one of the principal results of this case-study is the observation of a rather weak correlation between citation and topical relatedness. Thus, an attempt at projecting into the future the "topical relevance", in the limited sense of topic matching, from actual citations would in all likelihood be doomed to fail. The importance of a contextual citation analysis for the possible projection of future relevance relationships, it would seem to this author, is rather that of raising questions about what kind of use a document could be put to. By the description of the diverse functions served by citations to a document, we are naturally led to pose questions like: For what kind of research could the results of this document serve as evidence (ev)? What further research would be motivated (mo) in order to test the implications of this article? For what other purposes could the method (me) used in this paper be applied? Etc., etc.
7 Summary and conclusions

We set out in the beginning of this thesis to show or to make probable two things:
(1) that there are at least some cases, in which topic matching relationships are not necessary for explaining relevance judgments, as manifested by citations to articles in scientific journals, and
(2) that theoretical approach appears to influence relevance judgments, as manifested by citations to scientific papers, at least in some cases.

The methodology employed in the case-study performed in order to accomplish this was a contextual citation analysis, describing the functions of citations to two scientific articles from the research field of stratospheric ozone monitoring. The method involved three principal steps:
(i). Selection from scientific journals of a pair of articles, representing different theoretical approaches, but between which a topic matching relationship is holding, the criterion for topic matching being specified by the requirement of having at least two descriptors in common in a Dialog database. Further it was required that the two selected articles be as close as possible in time and location the influence on citation behavior of factors such as differences in exposure, impact and prestige of the journals in which the articles appeared.
(ii). Finding, by means of the SciSearch database, documents citing the selected articles and obtaining a fair portion of them (for documents citing one of the two originally selected articles a further selection was made by limiting the search to specific publication years).
(iii). Analysis of the obtained documents in order to determine the
(a) functions served by the citations to the initial pair of articles
(b) theoretical approach of citing articles, and how it matches with that of the cited source.
(c) topical relatedness between citing article and the initial pair of articles, both in absolute terms, in accordance with the criterion for topic matching, specified in (i) above, and as expressed by the union overlap measure (or similarity score), defined as the quotient between the number of shared descriptors and the total number of unique descriptors assigned in the database to a pair of documents.

Before entering into the case-study, however, an overview of previous research on relevance was given in chapter two, followed by a more detailed analysis of two concepts closely connected with relevance, viz. topicality and non-relevance. Chapter three dealt with citation analysis in a similar vein, giving a brief historical overview, followed by a more detailed discussion of citation functions and of some of the reasons for not citing. In chapter four the relationship between relevance judgments and citations was analyzed more closely, after which the methodology used in the case study was presented more in detail. In chapter five, finally, the results of the case-study, were presented.
In accordance with (i) above two articles from the field of research concerned with stratospheric ozone monitoring were selected, both published in 1974 in two prestigious multidisciplinary, scientific journals. The articles selected were seen to fulfill the criterion of topic matching, sharing two descriptors, as indexed in *Aerospace* database (Dialog file no. 108).

In accordance with (ii), 68 articles citing either or both of the two originally selected articles were acquired for analysis, in accordance with (iii), through consecutive steps of selection. The analysis involved the classification of each citation with respect to the function(s) it seemed to serve in the context of the citing article. This was effected by means of a classification scheme consisting of 14 diverse categories, that were developed mainly from a direct confrontation with the material analyzed. The results of the analysis were presented in tables of the distribution of citations over theoretical approach, citation functions and topical relatedness.

The results seem to indicate, among other things, that the correlation between citations and topical relatedness is rather weak, and that in particular there are several "anomalies" in the material studied, i.e. instances where the topical relatedness in terms of number of shared descriptors or similarity score, is actually lower for a cited than for a non-cited document. This can possibly be seen as a corroboration of (1) above.

The analysis undertaken in the case-study further showed an uneven distribution of the citations received by the two originally selected sources as for theoretical approach of the citing articles. One possible interpretation would be that the citing article is more likely to share the theoretical approach of the cited source, than to represent an opposite or alternative approach. This would seem to support (2) above.

Thus, as already noted, judging from the results of the case-study the correlation between citations and topical relatedness, as measured by the criterion for topic matching used here or by the union overlap measure, appears to be rather weak. Several interpretations of this result are possible. The most obvious observation to be made is perhaps, that the two articles selected as cited sources are already from the start topically related, fulfilling the requirement for topic matching between themselves and belonging to the same field of research. Hence, for any document that is topically related to one of these two originally selected articles, the likelihood that it will be topically related also to the other article, whether cited or not, is rather high; in any event it is definitely higher, than if the two articles had been randomly selected from totally different subject areas. Thus, the act of citation appears to be of less importance for topical relatedness, than the act of selection. But one of the points of this thesis was precisely to demonstrate the non-necessity of topic matching as an explanation for citations, and possibly also for relevance relationships. If the act of citation proves to be of lesser importance for topical relatedness, even for documents that are already topically linked, this could be seen as an indication that other factors than topical
relatedness are necessary in order to account for citations.\textsuperscript{43} It is the belief of this author that among these factors are the various \textit{functions} served by citations; functions that seem to endow them with some kind of relevance.

On a more general level it is, of course, possible to object, that the present material is too small and the method used is flawed, so that no conclusions whatsoever can be drawn. Another, more constructive way of looking at the matter would perhaps be too examine further, in a more comprehensive study, the relationship between topical relatedness and citations with respect to their ability to describe what relevance is. Assuming for a moment that there is indeed no obvious correlation between citations and topical relatedness, it would seem that there are two possibilities:

1. To reject citation functions as descriptive or predictive of relevance relations, and continue to build retrieval systems on topic matching or other types of topical relatedness.

2. To regard topic matching and related forms of measuring similarity between documents and queries as insufficient for describing relevance relationships, as manifested by the actual use of information through citations, and continue to develop methods for describing the relevance of citations by analyzing their functions, including also their location in the citing text (cf. Cano 1989).

One way of coming to a decision between these two alternatives would perhaps be to test the ability of both citations and topic matching to predict relevance judgments of scientists and scholars by employing methods similar to those of Cano (1989), that is using some kind of utility measure. That must, however, be the object of another study.

\textsuperscript{43} To state the argument in formal terms: let "C $\rightarrow$ T" mean that citation is important to topic matching, then "Not(C $\rightarrow$ T)" means that citation is not important to topic matching. But "C $\rightarrow$ T" is equivalent to "Not-T $\rightarrow$ Not-C", meaning that topic matching is \textit{necessary} to citation. Consequently, "Not(C $\rightarrow$ T)" is equivalent to "Not(Not-T $\rightarrow$ Not-C)", meaning that topic matching is not necessary to citation. Q.E.D.
References:


Duncan, E.B. et al. (1981), Qualified citation indexing: its relevance to educational technology. // Information retrieval in educational technology. - Aberdeen: University of Aberdeen. - P. 70-79


Foskett, D.J. (1972), A note on the concept of "relevance". // Information Storage and Retrieval. - Vol. 8, pp. 77-78


Garfield, E. (1979), Citation Indexing: its theory and application in science, technology, and humanities. - New York: John Wiley & Sons


Hill, W. J. & Sheldon, P. N.(1975), Statistical modeling of total ozone measurements with an example using data from Arosa, Switzerland. // *Geophysical research letters*. - Vol. 2: no. 12, pp. 541-544


Kamke, E. (1950), Theory of sets - New York: Dover


Lipetz, Ben-Ami (1965), Improvement of the selectivity of citation indexes to science literature through inclusion of citation relationship indicators. // American documentation. - Vol. 16: no. 2, pp. 81-90


90


Spiegel-Rössing, I. (1979), Science studies : bibliometric and content analysis. // Social studies of science. - Vol. 7: no. 1, pp. 97-113


93


94

Wilson, P. (1973), Situational relevance. // Information Storage and Retrieval. - Vol. 9, pp. 457-471


Zerefos, C.S., Repapis, C.C., Jenne, R.L (1982), Representatives of total ozone trends as derived from satellite BUV and ground-based measurements. // Pure and applied geophysics. - Vol.120: no.1, pp.29-53
Högskolan i Borås är en nationell högskola. Studenter kommer från hela landet för att studera här. Högskolan består av fem institutioner och utbildningsprogram och kurser ges inom områdena bibliotek och information, textil, teknik, pedagogik samt data och ekonomi.

Högskolans lokaler ligger mitt i centrala Borås.

1977 grundades Högskolan i Borås, men de textila utbildningarna har sitt ursprung ännu längre tillbaka i den statliga Tekniska Väfskolan som inrättades redan 1866.

Forskning och utvecklingsarbete är en expanderande del av högskolans verksamhet. Högskolan samverkar hår med företag, statliga myndigheter och kommuner i Sjuhäradsbygden och i andra delar av landet och utlandet.

Bibliotekshögskolan har funnits i Borås sedan 1972.

I många år var bibliotekshögskolan i Borås landets enda specialhögskola för bibliotekarier, men genom universitets och högskolornas nya frihet har bibliotekarieutbildningar inträtt även på andra universitet.

Bibliotekarieutbildningen leder till en magisterexamen som omfattar sammanlagt 160 poäng ( motsvarar åtta terminer), varav minst 80 poäng inom ämnet biblioteks- och informationsvetenskap.

Forskarutbildning i biblioteks- och informationsvetenskap ges vid Göteborgs universitet.

Bibliotek & läranderesurser (BLR) vid Högskolan i Borås har utfört digitaliseringen och har använt de exemplar som funnits i bibliotekets samlingar i befintligt skick.

De digitaliserade magisteruppsatserna är publicerade i Borås Akademiska Digitala Arkiv (BADA), som är högskolans system för digital publicering. http://bada hb.se

Uphovsrätten tillhör författarna.

Publiceringsår i BADA: 2014