

# INDEXING AS PROBLEM SOLVING: A COGNITIVE APPROACH TO CONSISTENCY

C. David and L. Giroux

Communication Department

Université de Montréal

P.O. 6128, Station A

Montréal, Québec H3C 3J7

S. Bertrand-Gastaldy and D. Lanteigne

École de biblioéconomie et des sciences de l'information

Université de Montréal

CP 6128, succ. Centre Ville

Montréal, Québec H3C 3J7

## **Abstract**

*Indexers differ in their judgment as to which terms reflect adequately the content of a document. Studies on interindexers' consistency identified several factors associated with low consistency, but failed to provide a comprehensive model of this phenomenon. Our research applies theories and methods from cognitive psychology to the study of indexing behavior. From a theoretical standpoint, indexing is considered as a problem solving situation. To access to the cognitive processes of indexers, three kinds of verbal reports are used. We will present results of an experiment in which four experienced indexers indexed the same documents. It will be shown that the three kinds of verbal reports provide complementary data on strategic behavior, and that it is of prime importance to consider the indexing task as an ill-defined problem, where the solution is partly defined by the indexer him(her)self.*

## **Résumé**

*Afin de mieux cerner l'origine des différences relevées par les études sur la cohérence inter-indexeurs, nous adjoignons les théories et méthodes issues de la psychologie cognitive (spécifiquement celles qui considèrent une tâche comme une situation de résolution de problème) à l'étude des processus*

*cognitifs impliqués lors de l'indexation. Cette approche permet d'identifier les stratégies cognitives élaborées et utilisées par les indexeurs. Afin d'avoir accès à leurs processus cognitifs, trois types de verbalisations sont recueillis. Nous présenterons les résultats d'une expérimentation pour laquelle quatre indexeurs expérimentés ont analysé les mêmes documents. Les résultats avancés démontreront la complémentarité des données issues des trois types de verbalisations et l'importance de considérer l'indexation en tant que problème mal- défini; la solution étant définie en partie par l'indexeur.*

## **Introduction**

Indexing is a costly process. Consequently, one needs some assurance with respect to product quality. The criteria of pertinence, exhaustivity and specificity help indexers in evaluating the validity of their work, but they are often difficult to operationalize (Rolling 1981; Bertrand-Gastaldy 1986). The interindexer consistency measure does, however, give an indication of the reliability of indexing, which is a partial demonstration of validity. A review of the literature reveals a rather weak level of interindexer consistency whatever the context (Leonard 1977; Markey 1984). While several studies of interindexer consistency have yielded a list of semantic, pragmatic and environmental factors having an influence on consistency (Zunde and Dexter 1969), all these studies share a concern for variations in the selected terms rather than for the nature of these terms or the processes which produce them. A model of the intellectual procedures involved in indexing remains to be elaborated. Indeed, many researchers have underlined the necessity of such a model and are actively working to fill this theoretical void (Belkin 1984; Begthol 1986; Endres-Niggemeyer 1990; Farrow 1991; Bertrand et al. 1994).

We argue that in order to understand the low level of interindexer consistency, one has to borrow from the theories and methods of cognitive psychology, and from problem solving theory in particular.

## **Problem Solving and Indexing**

Rather than viewing indexing as a routine cognitive situation or as text comprehension (Begthol 1986; Farrow 1991; Bertrand 1993), we will consider indexing as a problem solving activity, where one has to determine the themes contained in a document and produce a list of descriptors taken from a thesaurus. Subjects begin in a state of initial knowledge and move through the problem space to a final state or solution (Newell and Simon 1972). The problem space is the subject's representation of the task, that is, the possible states of the problem, the operations (cognitive or physical) that allow passage from one state to the next and all knowledge elements that the subject may judge useful to solve the problem. In order to move forward through the resolution process, indexers must, given the current problem state, select the appropriate operations. This selection activity is not aleatory; it depends on a control structure constructed by indexers as a function of their declarative and procedural knowledge as well as their appreciation of the attained state. In most cases, these operations can be grouped into sequences aimed at attaining a sub-goal, and the problem can be represented at a more general level as a structure consisting of goals and sub-goals, the nature and conditions of accomplishment of which are controlled by the subject. It is thus a matter of identifying and describing the components of an information processing system and the representations employed by this system in resolving a given problem (Newell and Simon 1972). In analyzing indexing as problem solving, it is necessary to identify the problem space specific to this activity. Following an examination of the task description in terms of norms and procedures (ERIC 1980; ISO 1985), we propose a hypothetical general problem space consisting of two large sets: a knowledge space and a resolution space (see Table 1). The knowledge space includes the set of declarative and procedural knowledge which are potential components of the problem. The resolution space consists of the major stages as defined by the norms and what we know of the usual indexing procedure. The bottom row of Table 1. shows examples of the form of knowledge used by an expert indexer familiar with a particular domain, a particular thesaurus and a given working environment.

(Footnote: Since indexing is a continuous process, it cannot be said that there is only one type of knowledge used in a single given step. However, certain steps may draw more on a particular knowledge set.) At the "scanning of document" stage, the expert's knowledge of indexing could be observed by a rapid recognition of the document's nature (a research report for example) and the inference of its structure after a few moments of scanning. For the content analysis" stage, the domain knowledge will serve to rapidly extract a rich macrostructure. At the concept selection" stage, knowledge of the domain and its macrostructure help in selecting the most central candidate concepts and simultaneously to evaluate their indexability and relevance as a function of terms in the thesaurus. It seems likely (Bertrand 1993; David 1990) that terms from the thesaurus encountered during the analysis are more easily considered by an expert. The "translation into descriptors" phase is shortened by the preceding identification, and the expert indexer's knowledge of the thesaurus makes for a precise and systematic consultation of the descriptors. During "revision," knowledge of elements of the context helps indexers to respect current policies in their work environment, to determine whether the selected descriptors meet the needs of the user clients of their respective institutions, and to ensure that the concepts represent or distinguish the content of the document in question relative to other similar documents in the database.

Our general working hypothesis thus supposes that in observing the strategic use of knowledge elements, it should be possible to explain variations in the choices of terms. With experts, for example, it can be assumed that with experience comes the assimilation of general rules and norms, knowledge relative to the particular work context, knowledge of the domain, etc., all of which leads to the development of a goal structure and sub-goals different from that of other indexers.

However, our quest is an inductive one and consists less in demonstrating a specific hypothesis than in identifying and modelling certain critical moments of the problem solving process of indexing.

---

**Table 1. General indexing problem space.**

Knowledge space

Resolution space

scanning of document

content analysis

concept selection

translation into descriptors

revision

Examples of expert knowledge usage

Indexing	Domain	Domain	Thesaurus	Context
knowledge	knowledge	knowledge	knowledge	knowledge

---

## Methodology

In order to shed light on the cognitive processes likely to account for variations in selected index terms, we asked four experienced indexers to index two documents each. (Footnote: The findings discussed here are taken from a larger task assigned to four novice indexers and four expert ones, with each indexing four documents: two with the Envirodoq thesaurus (1990) used daily by the expert indexers, two with the Infoterra thesaurus (1990) with which the experts had no familiarity. Only the findings pertaining to experts working with the Envirodoq thesaurus are presented here. It should be noted that indexing performed with the Envirodoq thesaurus involves primary descriptors (DE1) and secondary ones (DE2).) They all have considerable experience working in the

same document center on the environment and were thus familiar with the same procedures, as well as with the "Envirodoq" thesaurus developed for the ministère de l'Environnement du Québec and which contains about 900 descriptors.

In order to have access to the largest possible set of elements of the indexers' cognitive behaviour, we relied on three complementary methodological approaches grounded in verbalization by the subject: analysis of concurrent verbal report (think aloud method), retrospective verbalization, and an interview involving peer evaluation.

Verbal protocols are a well-known data gathering method in cognitive psychology. The very definition of a problem supposes that in such a situation subjects do not merely function automatically but consciously construct a representation of the problem and elaborate solution strategies. In this regard, it is possible to access these strategies inasmuch as they are at least partly explicit and under the individual's control. The most frequently used procedure is that of concurrent verbal reports, in which subjects are directed to think aloud during the execution of the task, thus providing a direct access" to their thought processes (Newell and Simon 1972; Ericsson and Simon 1980, 1984). Taken by itself, concurrent verbalization has serious limits for the analyst, particularly whenever the task being studied requires a certain expertise and is subject to external constraints not immediately visible as the task is being accomplished. It is for this reason that we also drew upon retrospective reports of the task (Hoc and Leplat 1983; Caverni 1988) which consist in showing subjects the traces (in this case, a video recording) of their own behaviour while questioning them about their interpretation of the process. Researchers can thus gain access to an entire set of data, the validity of which can at least be partially evaluated against the raw data of the concurrent verbal protocol. They also receive a description of certain global strategies or parameters adopted by the subject which are rarely verbalized. The third source of data is the peer evaluation interview. This consisted in asking the subjects to comment on the lists of descriptors produced

by other indexers. This type of verbal report was intended to bring out the evaluation criteria employed by each indexer with respect to the quality of a given indexing operation.

[Table 2. Lists of Descriptors Produced by the Expert Indexers \(12K\)](#)

**Findings**

Table 2 presents lists produced by the four expert indexers using a thesaurus familiar to them in indexing one of the two documents. The document in question was entitled *"Study of Sites for Aerated Ponds."* As can be easily seen, there is a clear difference in the degree of exhaustivity of the chosen terms. While Expert 1 only attributed two primary descriptors (DE1) and no secondary ones (DE2), Expert 4 attributed a total of seventeen descriptors, four DE1 and thirteen DE2. On the other hand, certain descriptors (*"aerated ponds"*, *used water treatment"*) are employed by all the indexers and the descriptors *"evaluation"* and *"site"* appear on three of the four lists. As such, the indexers are to a large extent in agreement with respect to certain central concepts contained in the document, while they differ dramatically in the attribution of secondary notions (DE2). The level of interindexer consistency for this document is 45% when all descriptors are included, and of 66 % if only primary descriptors are considered. (Footnote: The calculation of consistency was done by using Rolling's formula (1981).)

At first glance, analysis of the descriptors shows agreement among the expert indexers with respect to the basic semantic content, but variations with respect to the degrees of specificity, exhaustivity and exactitude. The point here is to understand why it is that indexers with the same general profile (same training, same expertise in the domain, and familiarity with the thesaurus) vary to this degree in the production of a list of descriptors. Following a preliminary analysis of the data collected by the above described methods, we modified our basic hypothesis. Initially, we conducted a detailed examination of the verbal protocols in order to grasp each goal structure. This was then juxtaposed with subsequent explanations and justifications, and in this way we expected to shed some light on the cognitive strategies underlying the observed indexer behaviour. However,

during the course of the analysis, two things became obvious. Firstly, the review of the respective goal structures did not reveal the anticipated important differences in behaviour. It is thus not easy to provide a satisfactory explanation for the low level of consistency on the basis of goal structures. In this regard, a recent study (Bertrand 1993) identified three main indexing resolution strategies, which vary as function of indexer's expertise. Among others, the "expert" strategy of indexers familiar with the controlled vocabulary leads to a knowledge driven list of to-be-selected terms. Indeed, at the very first stages of indexing, our expert subjects mentioned looking for thesaurus descriptors appearing in the document. The second observation pertains to the nature of the explanations and justifications provided by the indexers, as well as their evaluation of the other lists during the interview. The critical moment in the process is the decision to accept or reject a descriptor. In other words, while indexers may follow the same procedures (shared goals and sub-goals) in scanning documents, their criteria for deciding what constitutes good indexing seem to vary. This last observation led us to redirect our analysis. In doing so, it is necessary to recall that there are various types of problems. While these typologies cannot be described here, let us simply consider two large classes: well-defined and ill-defined problems (Simon 1973; Gick 1986). For well-defined problems, the information necessary for their resolution is entirely contained in the statement of the problem itself. One follows an often unique method of resolution requiring the application of very strict rules and often algorithmic procedures. Problems in algebra and physics, games like tic-tac-toe and Master Mind are examples of well-defined problems. Ill-defined problems, on the other hand, are characterized by incomplete or fuzzy representations of one or more of their basic components (goal, initial state, possible operations and constraints; Holyoak 1990). In ill-defined problems, problem space is often broader and less formal. Managing a business, renovating an apartment, writing a script or preparing a conference paper are examples of ill-defined problems. In this typology, indexing can be characterized



as an ill-defined problem inasmuch as, despite certain established indexing procedures, the nature of the goal is defined by the indexer.

It would appear that the key factor that needs to be understood if we are to explain the low level of interindexer consistency is variation in the representation of the solution. Indexing expertise is characterized by a gradual adaptation and transformation of directives learned during training into a more personalized mode of usage. Each indexer establishes an order of priority for the application of these rules and procedures. This order of priority derives from the criteria favoured by the indexer in the selection of terms constituting a good list of descriptors. While our analysis has yet to be completed, a few examples will show why this approach seems promising.

Generally speaking, the evaluation of descriptor lists by each of the indexers during the interview revealed just how much they are familiar with good indexing standards. They all clearly underlined the strengths and weakness of each indexing result in terms of these norms and could often correctly infer the criteria behind each choice of descriptor. As such, they all agreed about which list of descriptors best respected the norms, namely, the list prepared by Expert 2 (number of terms and their DE1 and DE2 distribution, appropriateness of the terms, and depth of the analysis). This particular list is indeed more respectful of indexing norms than the others: a good" list of descriptors should contain between eight and eleven descriptors, with DE1 containing the central notions and DE2 clarifying the latter, etc (ERIC 1980; ISO 1985; AFNOR 1985).

Retrospectively, the indexers can also consider their own work in light of this standard indexing scheme and discuss the criteria that lead them to differ from this standard.

In this sense, it is worth mentioning some aspects of their personal representation of the solution. This permits a partial explanation of differences in the lists. For Expert 1, a good" list of descriptors should represent the heart of a document and avoid creating false drop in retrieval. This accounts for his very limited choice of terms, *"when faced with the choice of including something about*

*which I am uncertain, I prefer putting nothing rather than creating false drop.*" For his part, Expert 4 does not index with particular users in mind, but with a concern for not omitting any notion which would subsequently go unnoticed. These criteria help in understanding the imbalance between his DE1 and DE2 lists and why he drifted away from the application of more standard criteria, *"in order to include everything in the documentt"*. While Expert 3's concern is with making sense for the user, he does not lose sight of the identification of primary notions, leaving aside those which strike him as secondary.

To better understand the dependency between an indexer's representation of the solution and his/her choice of descriptors, let us examine Expert 3's list in greater detail. In order to describe the main themes of document, Expert 3 is not satisfied with assigning disjointed descriptors. He attempts to assign them a role in a predicative structure. The first trace of this strategy is visible in Expert 3's concurrent verbalization when he attempts to compose a sentence with the descriptors: *"my document deals with sites for aerated ponds. For the treatment of used waters..."* The results of this strategy (his list of descriptors and, particularly, the order assigned to them) were criticized by Expert 4 and Expert 2, who judged the position assigned to *"evaluation"* and *"used water treatment"* to be inappropriate. In their view, *"evaluation"* should not appear at the top of the list because of its overly generic nature. *"Used water treatment"* should be in the DE1 list because it is a generic representation of the document's main subject matter and not a secondary aspect. A more precise explanation of these choices can be found in Expert 3's retrospective verbalization and interview. He states that the strategy underlying his choices is that of distinguishing those terms useful to retrieve pertinent analytical records from those which will be used afterward to distinguish records from one another. The term *"evaluation"* helps to distinguish the analytical record from the set of similar ones treating the same subject after the user has retrieved them. It should be noted that none of the other indexers evoke these criteria for the production of a list of descriptors. Expert 3 can be characterized as having a concern for the user and, as such, as

breaking certain rules (eg., avoiding the use of overly broad terms). Expert 3 appears to have created certain rules of his own, which are not always applied, but which distinguish him from the other indexers and allows for a cognitive explanation of his particular choice of descriptors. While there remains much more to analyze in our research, it is clear that the "problem solving" approach can explain aspects of the intellectual process of indexing which have received little attention in the literature on the subject. Moreover, the methods we have used appear to be a good way of bringing together the various traces of this process.

### **Acknowledgements**

This research was supported by a grant from the Social Sciences and Humanities Research Council of Canada and a grant from the Ministère des Affaires Internationales under the auspices of the France-Québec Cooperation Agreement.

### **References**

- Association Française de Normalisation. 1985. *Principes généraux pour l'indexation des documents*. AFNOR NF Z : 47- 102.
- Beghtol, Clare. 1986. Bibliographic classification theory and text linguistics: aboutness analysis, intertextuality and the cognitive act of classifying documents. *Journal of documentation*. 42, 2, june: 84-113.
- Belkin, N.J. 1984. Cognitive models and information transfer. *Social Science Information Studies* 4 : 111-29.
- Bertrand, A., Cellier, J.M., Giroux, L.1994. Expertise et stratégies dans une activité d indexation de documents scientifiques. *Le travail humain* 57, 1: 25-51.
- Bertrand, Annick.1993. *Compréhension et catégorisation dans une activité complexe. L'indexation de documents scientifiques*. Thèse de doctorat nouveau régime en psychologie. France: Université de Toulouse-Le Mirail.
- Bertrand-Gastaldy, Suzanne. 1986. De quelques éléments à considérer avant de choisir un niveau d'analyse ou un langage documentaire. *Documentation et bibliothèques* janvier-juin: 3-23.

- Bertrand-Gastaldy, S., Giroux, L., Lanteigne D., David, C. 1994. Les produits et processus cognitifs de l'indexation humaine. *ICO Québec, La gestion de l'information textuelle* 6, 1 et 2: 29-40.
- Caverni, Jean-Paul. 1988. La verbalisation comme source d'observables pour l'étude du fonctionnement cognitif. In Caverni, J.P., Bastien, C., Mendelsohn, P., Tiberghien, G., *Psychologie cognitive, modèles et méthodes*. Presse Universitaire de Grenoble: 253- 73.
- Centre de documentation et de renseignements. 1990. *Thésaurus Envirodoq*. Ministère de l'Environnement du Québec.
- David, Claire. 1990. *Élaboration d'une méthodologie d'analyse des processus cognitifs dans l'indexation documentaire*. Mémoire de maîtrise. Montréal: Université de Montréal.
- Endres-Niggemeyer, Brigitte. 1990. A procedural model of abstracting and some ideas for its implementation. *TKE'90, Terminology and Knowledge Engineering*. vol. 1. Proceedings of the Second International Congress on Terminology and Knowledge Engineering. INDEKS Verlag Frankfurt: 230-43.
- Educational Ressources Information Center. 1980. *ERIC Processing manual*, Section 7: Indexing, National Institute of Education: 409- 64.
- Ericsson, K.A., and Simon, H.A. 1984. *Protocol analysis: Verbal report as data*. Cambridge, MA: MIT Press.
- Ericsson, K.A., and Simon, H.A. 1980. Verbal Reports as Data. *Psychological review* 87, 3, may: 215-51.
- Farrow, John F. 1991. A cognitive process model of document indexing. *Journal of documentation* 47, 2, june: 149-166.
- Gick, M.L. 1986. Problem solving strategies. *Educational Psychologist* 21, 1&2: 99-120.
- Hoc, Jean-Michel, and Leplat, Jacques. 1983. Evaluation of different modalities of verbalization in a sorting task. *International Journal of Man-Machine Studies* 18: 283-306.

Holyoak, Keith. J. 1990. Problem solving. In Osherson, D., and Smith, E.E., ed., *Thinking: an invitation to cognitive science*. 3. Cambridge, MA: MIT Press: 117-46.

Leonard, Lawrence. E. 1977. *Inter-indexer consistency studies, 1954-1975: a review of the literature and summary of the study results*. Occasional papers. Graduate School of Library Science: University of Illinois.

MacMillan, J.T., and Welt, I.D. 1961. A study of indexing procedures in a limited area of the medical sciences. *American documentation* 12: 27-31.

Markey, Karen. 1984 Interindexer consistency tests: a literature review and report of a test of consistency in indexing visual materials. *Library and Information Science Research* 6: 155-77.

Newell, A.C., and Simon, H.A. 1972. *Human problem solving*. Englewood Cliffs, N.J.: Prentice-Hall.

Organisation internationale de normalisation. 1985. *Méthodes pour l'analyse des documents, la détermination de leur contenu et la sélection des termes d'indexation*. Norme Internationale ISO 5963.

Preschel, B.M. 1972. *Indexer consistency in perception of concept and in choice of terminology*. New York: Columbia University.

Programme des Nations Unies pour l'Environnement. 1990. *Infoterra : Thésaurus des termes relatifs à l'environnement*. PNUE. Kenya: Nairobi.

Rolling, L. 1981. Indexing consistency, quality and efficiency. *Information processing & Management* 17: 69-76.

Simon, Herbert A. 1973. The structure of Ill-Structured problems. *Artificial Intelligence* 4: 181-201.

Zunde, P., and Dexter, M.E. 1969. Factors affecting indexing performance. *Proceedings of the 32nd annual meeting of the American Society for Information Science*. ASIS. Greenwood publishing. CT. Westport : 313-22.