

# On conceptual models for information seeking and retrieval research

[Kalervo Järvelin](#)

Centre for Advanced Study  
University of Tampere, Tampere, Finland  
and

[T.D. Wilson](#)

Visiting Professor  
Högskolan i Borås, Borås, Sweden

---

## Abstract

There are several kinds of conceptual models for information seeking and retrieval (IS&R). The paper suggests that some models are of a summary type and others more analytic. Such models serve different research purposes. The purpose of this paper is to discuss the functions of conceptual models in scientific research, in IS&R research in particular. What kind of models are there and in what ways may they help the investigators? What kinds of models are needed for various purposes? In particular, we are looking for models that provide guidance in setting research questions, and formulation of hypotheses. As an example, the paper discusses [at length] one analytical model of task-based information seeking and its contribution to the development of the research area.

---

## Introduction

There has been considerable interest in recent years in producing conceptual models for information seeking and retrieval (IS&R) research. The recent paper by Wilson ([1999](#)) reviews models for information behaviour ([Wilson 1981](#)), information seeking behaviour ([Wilson 1981](#); [1996](#); [Dervin, 1986](#); [Ellis et al. 1993](#), [Kuhlthau, 1991](#)), and information searching or retrieval ([Ingwersen, 1996](#); [Belkin, et al. 1995](#); [Spink, 1997](#)).

Wilson ([1999](#): 250) notes concerning the models of information behaviour, among others, that "*rarely do such models advance to the stage of specifying relationships among theoretical propositions: rather they are at a pre-theoretical stage, but may suggest relationships that might be fruitful to explore or test.*" Later he notes that,

"[t]he limitation of this kind of model, however, is that it does little more than provide a map of the area and draw attention to gaps in research: it provides no suggestion of causative factors in information behaviour and, consequently, it does not directly suggest hypotheses to be tested." ([1999](#): 251)

It seems, therefore, that there may be several kinds of conceptual models for IS&R and that, at least for some research purposes, we would need models that may suggest relationships that might be fruitful to explore and provide hypotheses to test. The purpose of this paper is to discuss the functions of conceptual models in scientific research, in IS&R research in particular. What kind of models are there and in what ways may they help the investigators? What kinds of models are needed for various purposes? In particular, we are looking for models that provide guidance in setting research questions, and formulating hypotheses.

In the following section we shall discuss the meaning and function of conceptual frameworks and principles for judging their merits in research. This extends Järvelin's ([1987](#)) discussion on criteria for assessing conceptual models for IS&R research. Section 3 analyses briefly some summary frameworks for the IS&R domain. This is followed by a discussion of analytic frameworks. In particular, the classifications suggested by Järvelin are presented and their use in generating fruitful research hypotheses is discussed. Jarvelin's suggestions led to empirical study ([Byström & Järvelin, 1995](#); [Byström, 1999](#)) and theoretical development ([Byström, 1999](#); [Vakkari & Kuokkanen, 1987](#); [Vakkari, 1999](#)), which analysed the relationships of task complexity and information seeking. The uses for the classifications in later research are briefly summarised. The paper ends with discussion and conclusions.

## Conceptual models and their uses

All research has an underlying model of the phenomena it investigates, be it tacitly assumed or explicit. Such models, called conceptual frameworks ([Engelbart, 1962](#)) or conceptual models, easily become topics of discussion and debate when a research area is in transition. Often two or more models are compared and debated. With an eye on advancing the research area, how should the models be assessed for their possible uses? In this section we discuss the function of conceptual frameworks and principles for judging their merits.

According to Engelbart, developing conceptual models means specifying the following:

- Essential objects or components of the system to be studied.
- The relationships of the objects that are recognised.
- What kinds of changes in the objects or their relationships affect the functioning of the system - and in what ways.
- Promising or fruitful goals and methods of research.

Conceptual models are broader and more fundamental than scientific theories in that they set the preconditions of theory formulation. In fact, they provide the conceptual and methodological tools for formulating hypotheses and theories. If they are also seen to represent schools of thought, chronological continuity, or principles, beliefs and values of the research community, they become paradigms. The conceptual model of a research area is always constructed - it does not simply lie somewhere waiting for someone to pick it up.

The literature of the Philosophy of Science provides discussions on the functions of scientific theories. According to Bunge ([1967](#)), scientific theories are needed (or used) for the following functions:

- Systematisation of knowledge by:
  - Integrating formerly separate parts of knowledge.
  - Generalising and explaining lower abstraction level knowledge (or observations, data) through higher level constructs.
  - Explanation of facts through systems of hypotheses which entail the facts.
  - Expanding knowledge by deducing new propositions based on selected starting points and collected information.
  - Improving the testability of hypotheses through the control context provided by systems of hypotheses.
- Guiding research by:
  - Pointing to fruitful problems.
  - Proposing the collection of data, which nobody would understand to collect without the theory.
  - Proposing totally new lines of research.
- Mapping a portion of reality by:
  - Representing or modelling the objects (and relationships) of that chunk instead of just summarising the data.
  - Providing a tool for producing new data.

We believe that these functions are also suitable functions of conceptual models, which are more general in nature than theories. Clearly, conceptual models may and should map reality, guide research and systematise knowledge, for example, by integration and by proposing systems of hypotheses.

A conceptual model provides a working strategy, a scheme containing general, major concepts and their

interrelations. It orients research towards specific sets of research questions. A conceptual model cannot be assessed directly empirically, because it forms the basis of formulating empirically testable research questions and hypotheses. It can only be assessed in terms of its instrumental and heuristic value. Typically, this happens by assessing the research strategies and programmes (and results) it creates. The latter programmes consist of interrelated substantial theories and research relevant for evaluating them ([Wagner, et al., 1992](#); [Vakkari 1998](#)). If the substantial theories prove to be fertile, the model is so too.

However, waiting for the substantial theories to prove to their fertility may take some time. In the meantime, or even before embarking on some line of research, it may be important to argue about the merits of various conceptual models. The following are the types of arguments that can be used to judge the merits of a conceptual model:

- General scientific principles:
  - When studying some phenomena, they should be studied in all situations, and also under extreme conditions (cf. thermophysics). Thus, you do not just consider information seeking by academics but also by other professions or by laymen.
  - The framework should be limited in a meaningful way as a system. For understanding information seeking by human actors, the proper system is not some service (for example, a library) and its clients but rather an information actor immersed in his or her situation and information environment (for example, all information access systems).

When two competing conceptual models are compared the following criteria may be applied to judge their merits:

- Simplicity: simpler is better other things being equal.
- Accuracy: accuracy and explicitness in concepts is desirable.
- Scope: a broader scope is better because it subsumes narrower ones, other things being equal.
- Systematic power: the ability to organise concepts, relationships and data in meaningful systematic ways is desirable.
- Explanatory power: the ability to explain and predict phenomena is desirable.
- Reliability: the ability, within the range of the model, to provide valid representations across the full range of possible situations.
- Validity: the ability to provide valid representations and findings is desirable.
- Fruitfulness: the ability to suggest problems for solving and hypotheses for testing is desirable.

Theoretical development or the construction of new conceptual models in any research area often requires conceptual and terminological development. Conceptual development may mean fulfilling, perhaps in a better way than before, the basic requirements for scientific concepts - precision, accuracy, simplicity, generality, and suitability for expressing propositions, which may be shown true or false. Moreover, good concepts represent essential features (objects, relationships, events) of the research area. More importantly, the concepts should differentiate and classify the phenomena in ways that lead to interesting hypotheses (or research problems). This means that the concepts must relate to each other in systematic and fruitful ways. Concepts also need to support research into the phenomena by known research methods (or, somewhat relaxed, by methods that can be developed). They need to be compatible with each other and with research methods (that is, be congruent).

## Summary frameworks

### Two sample frameworks

We will discuss Ellis's ([1989](#); [Ellis, et al., 1993](#)) and Ingwersen's ([1996](#)) frameworks. These are used and discussed here as examples only and we make no claims about their merits with respect to the research tasks for which they were originally intended.

Ellis's elaboration of the different behaviours involved in information seeking consists of six features. Ellis makes no claims to the effect that the different behaviours constitute a single set of stages; indeed, he uses the term 'features' rather than 'stages'. These features are named and defined below:

- Starting: the means employed by the user to begin seeking information, for example, asking some knowledgeable colleague;

- Chaining: following footnotes and citations in known material or 'forward' chaining from known items through citation indexes;
- Browsing: 'semi-directed or semi-structured searching' (Ellis, 1989: 187);
- Differentiating: using known differences in information sources as a way of filtering the amount of information obtained;
- Monitoring: keeping up-to-date or current awareness searching;
- Extracting: selectively identifying relevant material in an information source;
- Verifying: checking the accuracy of information;
- Ending: which may be defined as 'tying up loose ends' through a final search.

The strength of Ellis's model is that it is based on empirical research and has been tested in subsequent studies, most recently in the context of an engineering company ([Ellis & Haugan, 1997](#)).

Of the features, Ellis ([1989](#): 178) notes that, '*...the detailed interrelation or interaction of the features in any individual information seeking pattern will depend on the unique circumstances of the information seeking activities of the person concerned at that particular point in time*'. Wilson (1999) proposes how these features may relate to each other temporally, providing a partial order; see Figure 1.

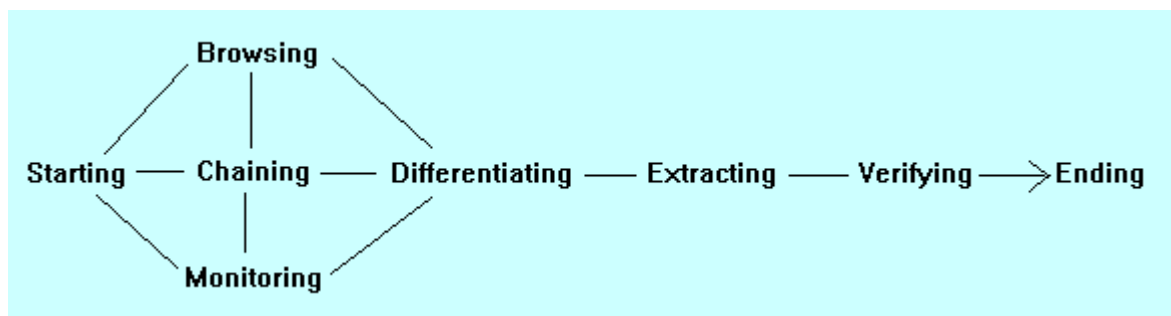


Figure 1. A process version of Ellis's behavioural framework ([Wilson 1999](#))

One may *describe* any information seeking activities through Ellis's features. Indeed, they are general enough to fit a large number of empirical situations. However, if one is to *explain* information seeking behaviour, say, in terms of the work tasks the subjects are engaged with, or their knowledge on the task, the features fall short because they are not explicitly related to such external possible causative factors.

Of course, Ellis's model may still be of indirect help in finding explanations for information seeking behaviour. It is possible to discern differences in any of the 'features' in different situations, involving different kinds of persons through successive research projects. For example, some persons in some roles may be shown to engage more or less in monitoring than other persons. This may then lead to an examination of the factors that 'cause' these differences.

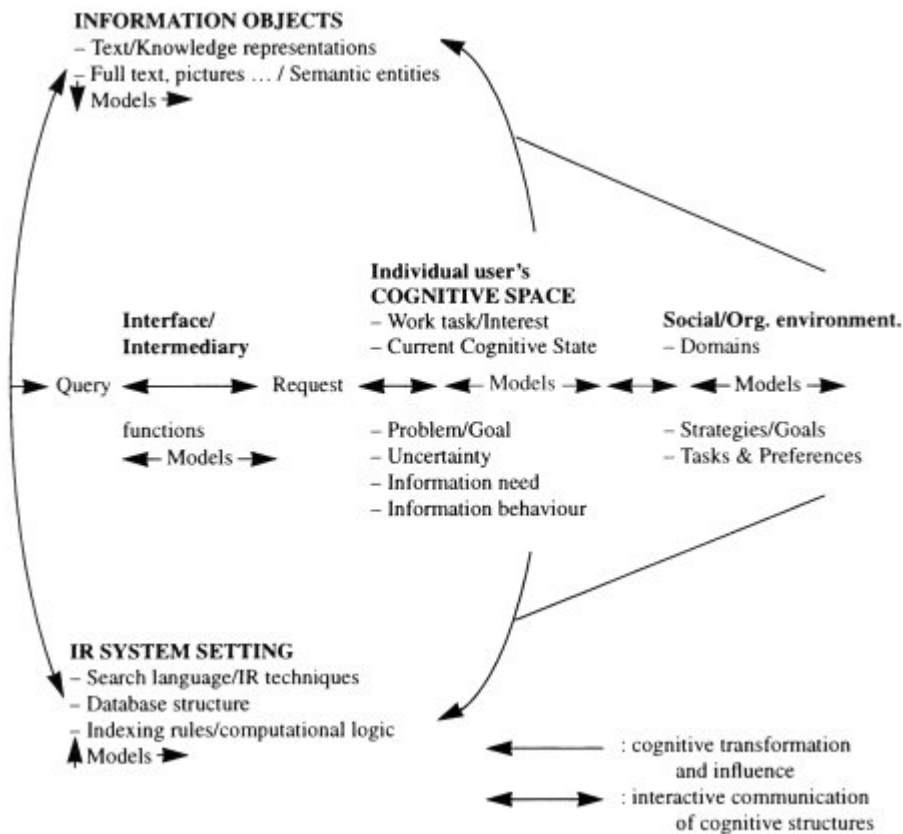


Figure 2. Ingwersen's model of the IR process (Wilson, 1999; based on [Ingwersen, 1996](#))

Ingwersen's ([1996](#)) model is slightly simplified in Figure 2. Wilson points out its relationships to other models of information seeking behaviour. In particular, the elements *user's cognitive space* and *social/organisational environment*, resemble the *person in context* and *environmental factors* specified in Wilson's models ([1981](#), [1996](#); [1999](#)). The general orientation towards queries posed to an IR system point to a concern with the *active search*, which is the concern of most information-seeking models. Ingwersen, however, makes explicit a number of other elements: first, he demonstrates that within each area of his model, the functions of the information user, the document author, the intermediary, the interface and the IR system are the result of explicit or implicit cognitive models of the domain of interest at that particular point. Thus, users have models of their work-task or their information need, or their problem or goal, which are usually implicit, but often capable of explication. Again, the IR system is an explication of the system designer's cognitive model of what the system should do and how it should function. Secondly, Ingwersen brings the IR system into the picture, suggesting that a comprehensive model of information-seeking behaviour must include the system that points to the information objects that may be of interest to the enquirer. Thirdly, he shows that various cognitive transformations take place in moving from the life-world in which the user experiences a problem or identifies a goal to a situation in which a store of pointers to information objects can be satisfactorily searched and useful objects identified. Finally he points to the need for these cognitive structures and their transformations to be effectively communicated throughout the 'system', which will include the user, the author and the IR system designer. All this true—it is easy to agree.

Thus, Ingwersen's model, to a degree, integrates ideas relating to information behaviour and information needs with issues of IR system design, and this is an important strength of the model. Saracevic suggests that ([1996](#)): '*The weakness is in that it does not provide for testability... and even less for application to evaluation of IR systems.*' However, recently, Borlund and Ingwersen ([1997](#); [1998](#); [Borlund, 2000](#)) have developed and tested an evaluative strategy on the basis of this model and have demonstrated its value in testing interactive IR systems. A remaining potential weakness is that information behaviour other than information retrieval is not explicitly analysed. Issues of how users arrive at the point of making a search, and how their cognitive structures are affected by the processes of deciding how and when to move towards information searching, may be lost. These issues may be discussed in terms of the social or organisational environment but, to say the least, this is not explicit.

In Ingwersen's model, there are several entities of the IS&R interplay present, and some of their relevant features are explicated. Therefore, there are better possibilities for formulating research questions for empirical study; for example, how is an individual user's uncertainty related to the intermediary functions, and how does this affect the retrieval process? However, there is still some way to go before one may say that an empirical research problem has



been specified. This could be done by classifying, for example, uncertainty and intermediary functions in ways that suggest empirical relationships.

## Uses of summary frameworks

Summary models provide overviews of research domains, and list factors affecting the phenomena. It is often easy to agree that, what the models propose, are factors affecting the processes of interest. However, without detailed analysis of the components, such models provide little or no suggestion of causative factors in IS&R phenomena and, consequently, they do not directly suggest hypotheses to be tested. Indirectly, however, a comparison of findings across several studies may suggest causative factors to be explored.

## An analytic framework

Järvelin ([1987](#)) suggested three classifications and discussed their use in generating fruitful research hypotheses for the analysis of the relationships of task complexity and information seeking. Byström and Järvelin ([1995](#); [Byström 1999](#)) revised the classification and carried out an empirical study, and Byström ([1999](#)) and Vakkari ([1998](#); [1999](#)) suggested theoretical developments. We first present the classifications and then discuss their theoretical and methodological consequences.

## Task complexity

A worker's job consists of tasks, which consist of levels of progressively smaller subtasks. Tasks are either given to, or identified by, the worker. Each task has a recognisable beginning and end, the former containing recognisable stimuli and guidelines concerning goals and/or measures to be taken ([Hackman, 1969](#)). Seen in this way, both a large task or any of its (obviously simpler) sub-tasks may be considered as a task. This relativity in definition is necessary in order to analyse tasks of different levels of complexity.

In information seeking we are interested in information-related tasks. These can be seen as perceived (or subjective) tasks or objective tasks. The relationships of objective and perceived tasks have been considered in organisational psychology ([Campbell, 1988](#); [Hackman, 1969](#); [Wood, 1986](#)) where task descriptions based on perceived tasks are generally held invalid for many purposes (for example, [Roberts & Glick, 1981](#)). However, in information seeking, perceived tasks must be considered because each worker may interpret the same objective task differently (for example, as regards its complexity) and the perceived task always forms the basis for the actual performance of the task and for interpreting information needs and the choice of promising actions for satisfying them.

The literature suggests many task characteristics related to complexity: repetition, analysability, *a priori* determinability, the number of alternative paths of task performance, outcome novelty, number of goals and conflicting dependencies among them, uncertainties between performance and goals, number of inputs, cognitive and skill requirements, as well as the time-varying conditions of task performance ([Campbell, 1988](#); [Daft et al., 1988](#); [Fischer, 1979](#); [Fiske & Maddi, 1961](#); [Hart & Rice, 1991](#); [Järvelin, 1986](#); [March & Simon, 1967](#); [MacMullin & Taylor, 1984](#); [Tiamiyu, 1992](#); [Tushman, 1978](#); [Van de Ven & Ferry, 1980](#); [Wood, 1986](#); [Zeffane & Gul, 1993](#)). Also, these characteristics have been understood in many different ways in the literature. They belong in two main groups: characteristics related to the *a priori* determinability of tasks, and characteristics related to the extent of tasks.

Järvelin ([1987](#); [Byström and Järvelin, 1995](#)) suggest a simple, one-dimensional categorisation of the complexity of tasks based on, from the worker's point of view, *a priori* determinability of, or uncertainty about, task outcomes, process and information requirements. This dimension is related to the above task characteristics: repetition, analysability, *a priori* determinability, the number of alternative paths of task performance and outcome novelty. Similar one-dimensional categorisations of complexity are used by Tiamiyu ([1992](#)) and Van de Ven and Ferry ([1980](#)). Simple tasks are routine information processing tasks, where the inputs, process and outcomes can be determined *a priori*, while difficult or complex tasks are new and genuine decision tasks, which cannot be so determined. Such a categorisation is generic and, thus, widely applicable to many types of tasks and domains.

## Task categorisation

In this paper, tasks are classified into five categories ranging from an automatic information-processing task to a genuine decision task. This categorisation is based on the *a priori* determinability (or structuredness) of tasks and is closely related to task difficulty or complexity.

Task complexity is often seen to depend on the degree of *a priori* uncertainty about the task inputs, process and outcome (for example, [Van de Ven & Ferry, 1980](#)). In automatic information processing tasks, the type of the task result, the work process through the task, and the types of information used can all be described in detail in advance. In genuine decision tasks, on the contrary, none can be determined *a priori*.<sup>1</sup> Our task categorisation is presented in Fig. 3 where information (both input and result) is represented by arrows and the task process by boxes. The *a priori* determinable parts of tasks are represented by solid arrows and solid boxes, and the *a priori* indeterminable parts of tasks are represented by dashed arrows and shaded boxes. Dashed arrows and shaded boxes thus represent case-based arbitration. Three arrows are used in the input side to visualise that many inputs often are needed and that there are degrees of *a priori* determinability among them. Also the types of input differ by task category as discussed in the next subsection.

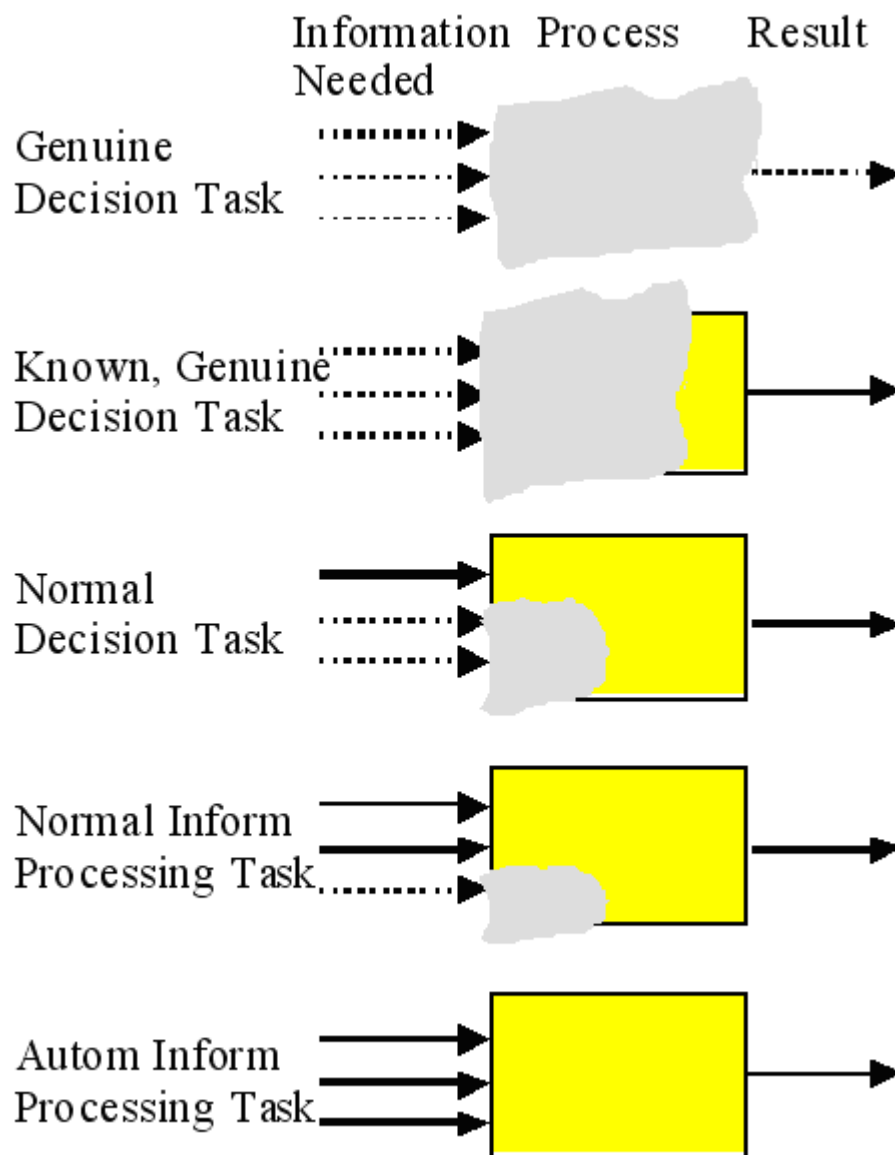


Figure 3: Task categories (Anon. 1974)

Tasks in different categories can be characterised briefly as follows:

- *Automatic information processing tasks* are *a priori* completely determinable so that, in principle, they could be automated - whether actually automated or not. Example: computation of a person's net salary yields a real number in some known range and requires this person's gross salary and tax code, and the taxation table.

*Normal information processing tasks* are almost completely *a priori* determinable but require some case-based arbitration concerning for example, the sufficiency of the information normally collected. Thus part of the process and information needed is *a priori* indeterminable. Example: tax coding is mostly rule-based but some cases require additional clarification, that is, case-dependent information collection.

- *Normal decision tasks* are still quite structured but in them case-based arbitration has a major role. Example: hiring an employee or evaluating a student's term paper.
- In *known, genuine decision tasks* the type and structure of the result is *a priori* known but permanent procedures for performing the tasks have not yet emerged. Thus, the process is largely indeterminable and so are its information requirements. Example: deciding about the location for a new factory or medium-range planning in organisations.
- *Genuine decision tasks* are unexpected, new and unstructured. Thus, neither the result, the process nor the information requirements can be characterised in advance. The first concern is task structuring. Example: the collapse of the Soviet Union from the viewpoint of other governments.

Information seeking research has focused mostly on tasks in the middle and upper parts of the categories (normal decision task to genuine decision task) although this dimension has only rarely been recognised. Belkin (1980) describes a similar scale of problem situation levels). The categories above are relative to the worker: what is a genuine decision task to a novice may be a normal decision to an expert.

## Types of information needed in tasks

In expert systems design, the types of information are classified as problem information (PI), domain information (DI), and problem solving information (PSI) (for example, Barr & Feigenbaum, 1981). Järvelin and Repo (1983; 1984) proposed these concepts for information seeking research. These information categories can be characterised as follows:

- *Problem information* describes the structure, properties and requirements of the problem at hand. For example, in bridge construction, information on the type and purpose of the bridge and on the building site constitute problem information. It is typically available in the problem environment, but, in the case of previous problems of the same type, it may also be available in documents.
- *Domain information* consists of known facts, concepts, laws and theories in the domain of the problem. For example, in bridge construction, information on the strength and thermal expansion of steel belongs to domain information. This is, typically, tested scientific and technological information published in journals and textbooks.
- *Problem-solving information* covers the methods of problem treatment. It describes how problems should be seen and formulated, what problem and domain information should be used (and how) in order to solve the problems. For example, in bridge construction, the design engineer's heuristics concerning the pros and cons of various bridge design types constitute problem-solving information. It is instrumental information and typically available only from knowledgeable persons (or experts).

These three information categories are orthogonal, that is, represent three different dimensions and have different roles in problem treatment. All are necessary in problem treatment but, depending on the task, and to different degrees, may be available to a worker performing the task. Because their typical sources are different, typical channels for acquiring them may also be different.

Regarding Figure 3, the solid arrows representing input information may be seen as *a priori* determinable problem information whereas the dashed arrows would represent all *a priori* indeterminable information, often problem-solving information.

## Types of Information Sources

Byström and Järvelin (1995) classified the types of information sources as:

- *fact-oriented*:
  - registers (manual and computerised catalogues and files)
  - commercial databases
- *problem-oriented*:



the people concerned (for example, people proposing, or affected by, administrative actions)

- official documents (for example, agendas, meeting minutes, letters, applications, memoranda, maps, unpublished planning documents)
- *general-purpose*:
  - experts (including knowledgeable colleagues)
  - literature (for example, books, reports, journals, newspapers)
  - personal collections (personal notes, calculations, etc.)

They also classified the sources as being either *internal* or *external* to the organisation in which the user works.

## Theoretical and methodological consequences

Byström and Järvelin used their framework, the three classifications of tasks, information and information sources, for the analysis of their data structured in work charts (Figure 4). In combination, the three classifications suggest a set of hypotheses of the type: "*Tasks of complexity type X require information of type Y that is available from sources of type Z*". Thus the classifications suggest analytical relationships between the variables.

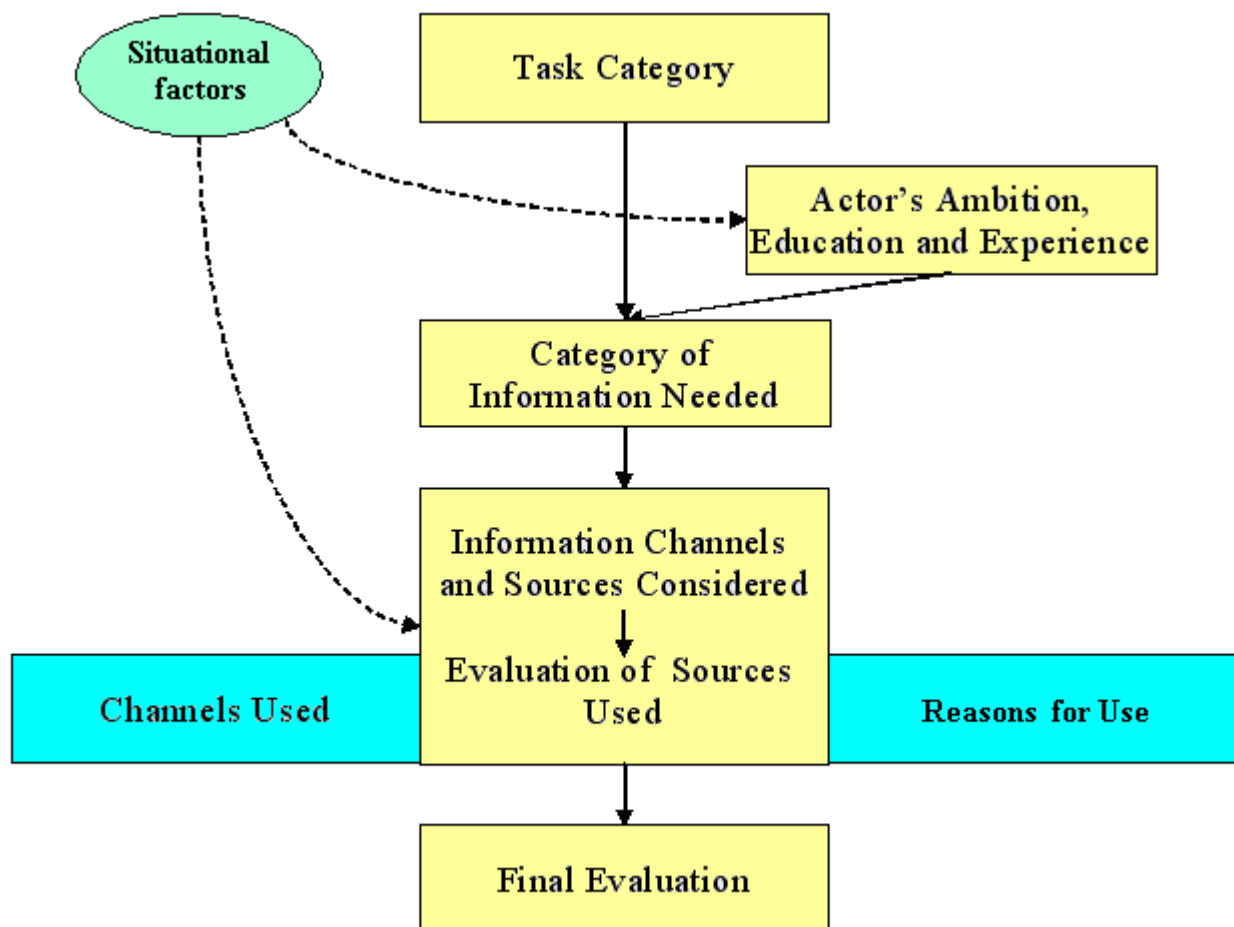


Figure 4: The work chart structure (Byström and Järvelin, 1995.)

Byström and Järvelin (1995; Murtonen, 1992<sup>2</sup>) developed a qualitative method for task-level analysis of the effects of task complexity on information seeking and found, in a public administration context, that these effects are systematic and logical. The specific research problem studied was: what types of information are sought through which types of channels from what kinds of sources in which kinds of tasks? They found that, as task complexity increased, so:

- the complexity of information needed increased,
- the needs for domain information and problem solving information increased,
- the share of general-purpose sources (experts, literature, personal collections) increased and that of problem and fact-oriented sources decreased,

- the success of information seeking decreased,
- the internality of channels decreased, and
- the number of sources increased.

The contrast between simple and complex tasks underlines the importance and consequences of task complexity: in the latter understanding, sense-making and problem formulation are essential and require different types and more complex information through somewhat different types of channels from different types of sources.

Byström followed on with further empirical studies ([1999](#); [Murtonen, 1994](#)). Based on her empirical findings, Byström presented a revised model of task-based information seeking (Figure 5). The model contains eleven statements (S1 - S11 in Figure 5). Some of the statements are given below (all are given in the [Appendix](#)):

- S2: the more information types are needed, the greater the share of people as sources.
- S6: the higher the degree of task complexity, the more probable is the need for multiple information types: first task information, then task and domain information, and finally task, domain and [problem] solving information.
- S8: the higher the degree of task complexity, the more information types are needed and the greater the share of general-purpose sources and the smaller the share of task-oriented sources.
- S10: task complexity is distinctly related to increasing internality of people as sources and decreasing internality of documentary sources.
- S11: Increasing task complexity fosters the use of people as sources.

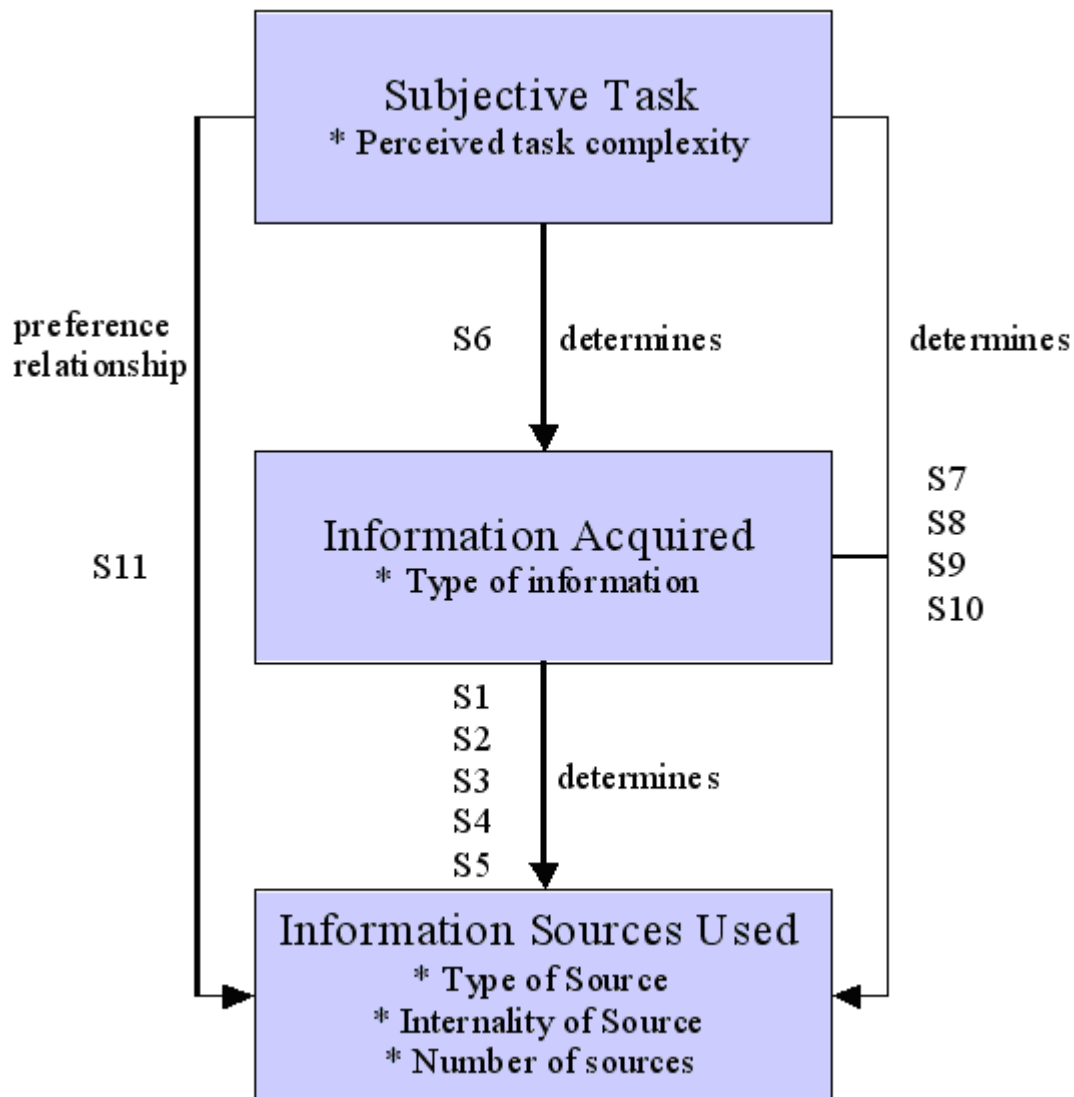


Figure 5: A model of task-based information seeking ([Byström, 1999](#))

Vakkari ([1998](#); [1999](#); [Vakkari & Kuokkanen, 1997](#)) analysed, and contributed to, theory growth in task-based

information seeking. Vakkari and Kuokkanen apply Wagner & Berger's (1985) analysis of theory growth to reconstructing a theory based on the framework by Byström and Järvelin (1995). Vakkari and Kuokkanen note that the latter did not fully utilise the whole potential of the framework, for example, the relationships of information types and source use was not fully developed. They derive new hypotheses for further empirical work from the reconstructed theory. The resulting theory is thus broader in scope and has more empirical consequences than the original. Vakkari and Kuokkanen state that their reconstruction creates potential growth of knowledge within the theory of information seeking. This is easy to agree.

Vakkari (1998) further uses Wagner & Berger (1985) and focuses on the theoretical research programme starting from Tushman's (1978) study on task complexity and information. He finds that Byström and Järvelin's (1995) work created progress in all dimensions of theory growth, especially in terms of precision and scope. The framework (research programme), by adding the classification of information types, explicated several new factual relations among information seeking phenomena.

The empirical findings and theoretical developments by Byström, Järvelin and Vakkari classify tasks, information and information sources in a systematic way. The latter are also systematically related to other central concepts of information seeking in a systematic way. The original papers suggested some classifications of essential phenomena. The original classifications were really simple, even trivial, when presented. However, they suggested specific systematic relationships to be explored. This led, in later papers, to thorough empirical work and theoretical development. This is an example of how proper analytic models may aid research in a specific area, such as information seeking.

## Discussion and conclusions

The previous section presented a framework for information seeking studies that directly suggested research questions and hypotheses for testing. Such frameworks are clearly needed in building up a knowledge base in the IS&R domain. Unfortunately, the work discussed above is not complete and we cannot present a well thought-out complete framework. There is room for further work, which is not the purpose of the present paper. Moreover, the model discussed is very specific, it does not attempt to cover all phenomena related to [task-based] information seeking.

However, as a small contribution to further development, we can point to the fact that the model makes no reference to the characteristics of the person (apart from the possibility that novices and experts will behave differently), or to the field in which the person works. Other investigations have drawn attention to individual personality as a determinant of information-seeking behaviour (e.g., Kernan & Mojena, 1973; Bellardo, 1985; Palmer, 1991), and to the discipline or context within which the person works (e.g., Anon., 1965; Auster & Choo, 1994; Fabritius, 1998; Greene & Loughridge, 1996; Herner & Herner, 1967; Siatri, 1998; Timko & Loynes, 1989; Wilson & Streatfield, 1980). For example, the fact that more complex decisions involve more searching for people as sources of information may differ depending upon the person's 'need for affiliation' (McClelland, 1961).

From the point of view of context or discipline, even in the field of public administration, for example, there may be significant differences in the nature of the tasks in, say, a planning department and a more 'people oriented' department such as social work. In the former, the processing of applications may involve much more decision making of a formal, technical nature, while in the latter, the concern with people's personal and domestic problems may result in decisions that have consequences that are more difficult to assess. We can suggest, therefore, a distinction between decisions that are related to a 'concern for process' and those that are related to a 'concern for person'.

We can also note that the distinction between 'information' and 'advice' is not sufficiently explored, although we suspect that the increased use of people as sources in complex decisions may have as much to do with the ability of people to guide, evaluate and advise, as with their possession of expert knowledge. Previous work on the affective dimension of information behaviour may also be relevant here (Wilson, 1981; Kuhlthau, 1993).

Finally, we can also point to a second dimension of decisions: as noted above, the present framework uses one dimension "a priori determinability of, or uncertainty about, task outcomes, process and information requirements". Thompson (1967) proposed two dimensions, one of which is similar to that used here, "Preference regarding possible outcomes", which might be 'certainty' or 'uncertainty'. The second dimension is "Beliefs about cause/effect

relationships", which, again, might be 'certain' or 'uncertain'. The matrix that results from the combination of these two dimensions gives four types of decision processes, as shown in Figure 6.

Beliefs about cause/effect relations	Preferences regarding possible outcomes	
	Certainty	Uncertainty
Certain	<i>Computation</i>	<i>Compromise</i>
Uncertain	<i>Judgement</i>	<i>Inspiration</i>

Figure 6: Decision processes (based on Thompson, 1967)

The conceptual richness that results from the addition of a second dimension would give rise to an additional set of hypotheses relating the use of information sources to decision process. For example, one might hypothesise that decisions requiring 'judgement' will involve more information seeking activity and a greater use of discussions with colleagues, than other types of decision process, while 'inspiration' may require more personal 'thinking time' and use of a greater variety of information sources.

We return to the requirements on conceptual frameworks presented above. The framework developed by Byström, Järvelin and Vakkari, through several studies, may be claimed to meet several of the requirements. In Engelbart's terms, it suggests that tasks, information, and information channels and sources are central objects in information seeking. It further suggests how these objects are related to each other. The hypotheses generated were (are) fruitful goals for further research.

Regarding Bunge's (1967) functions for scientific theories, here applied for assessing conceptual frameworks, we find the following when assessing the Byström, Järvelin and Vakkari framework:

- Systematisation of knowledge by:
  - Integrating formerly separate parts of knowledge: *Task complexity studies from organisational research are integrated with information seeking studies.* .
  - Generalising and explaining lower abstraction level knowledge (or observations, data) through higher level constructs: *Specific information needed and sought, much studied in information seeking, is analysed in terms of types of information.* .
  - Explanation of facts through systems of hypotheses, which entail the facts: *The framework suggested and allowed verification of several hypotheses of the research domain, cf. Byström's S1-S11.* .
  - Expanding knowledge by deducing new propositions based on selected starting points and collected information: *The later empirical and theoretical developments clearly expanded the original approach - Vakkari and Kuokkanen added to the original findings and unit theory.* .
  - Improving the testability of hypotheses through the control context provided by systems of hypotheses: *The classifications generated many related hypotheses (e.g., Byström's S1 - S11) which provided, for each hypothesis, a context for its verification.* .
- Guiding research by:
  - Pointing out fruitful problems: *From the beginning, the framework saw tasks, as opposed to whole jobs, related to information seeking through the types of information needed in tasks. The latter were seen to vary along task complexity.* .
  - Proposing the collection of data, which nobody would think to collect without the theory: *The framework suggested data to be collected on task complexity, task-related information seeking and the types of information needed. These were novel ideas in the late 1980s in information seeking research.*
  - Proposing totally new lines of research: *The framework was one approach, among others, towards the task-centred line of information seeking research.*
- Mapping an area of reality by:
  - Representing or modelling the objects (and relationships) of that area instead of just summarising the data: *While early research in information seeking summarised job-level information seeking, sources and preferences, the framework suggested tasks and information types as explaining the phenomena.* .
  - Providing a tool for producing new data: *The framework was a useful tool for generating hypotheses, and the associated research methods allowed the production of the required data.*

Regarding general scientific principles, suggested above, for the assessment of conceptual frameworks, we may point out the following:

- The framework is general in the sense that it supports the analysis of task-based information seeking for any kinds of tasks through categories that are not limited to special contexts, for example, academics. The tasks need not be job-related, leisure tasks do as well.
- The framework suggests perceived tasks, needed information and information seeking as a meaningful system. From the person or actor viewpoint this is much more meaningful than the information source and system framework (of many earlier studies) alone.

Further desiderata for conceptual models were:

- Simplicity: simpler is better other things being equal. *The framework is based on very simple classifications.*
- Accuracy: accuracy and explicitness in concepts is desirable. *The framework could be more accurate and explicit in its classification on task complexity. Nevertheless, it has functioned well as a first approximation. The framework is more accurate than its predecessors in its focus on task-level instead of job-level.*
- Scope: a broader scope is better because it subsumes narrower ones, other things being equal. *The framework is broader in its hospitability to any kind of tasks, not just job-related. On the other hand, it covers just three concepts, albeit important ones, of information seeking - a broader framework would incorporate other concepts as discussed above.*
- Systematic power: the ability to organise concepts, relationships and data in meaningful systematic ways is desirable. *This clearly is one strong feature of the framework.*
- Explanatory power: the ability to explain phenomena reliably and to predict them is desirable. *This clearly is one strong feature of the framework; it suggested several hypotheses that were later confirmed.*
- Validity: the ability to provide valid representations and findings is desirable. *(No model can directly argue for being valid)*
- Fruitfulness: the ability to suggest problems for solving and hypotheses for testing is desirable. *The number of studies that followed suggests at least some fruitfulness.*

We do not wish to make any claims about the usefulness or significance of this framework in comparison to other approaches within information seeking research. Rather, we wish to point out its formal merits: because of its characteristics, it has been successful in generating research that seems to have led to empirical and theoretical developments in the area of information seeking. Such models are needed in information science. According to Vakkari and Kuokkanen (1997), in order to create new knowledge in information science, we need clear, conceptually structured descriptions of the research objects. Without them the utilisation of research results in further studies is hampered. That would lead to slow or non-existent growth of knowledge in the field while findings may still amass.

---

Send your comments on this paper to the journal's discussion list - join [IR-discuss](#)

---

## Notes

1. It is this factor of determinability that helps us to define the 'automatic information processing' task. In such a task the outcome is determinable in advance. While a computer may be programmed to undertake tasks, which are computationally simple, such as those in a chess game, the outcome of the computer's calculations will not be determinable in advance, because of the complexity of the game.
2. Murtonen is the maiden name of Byström.

## References

- Anon. (1965). Survey of information needs of physicists and chemists. *Journal of Documentation*, **21**(2), 83-112.
- Anon. (1974). *Tietosysteemin rakentaminen* [Information system design]. Helsinki: Tietojenkäsittelyliitto. (Publication no. 25). (In Finnish).



- Auster, E., & Choo, C. W. (1994). How senior managers acquire and use information in environmental scanning. *Information Processing and Management*, **30**(5), 607-618.
- Barr, A. & Feigenbaum, E., (Eds.), (1981). *Handbook of artificial intelligence: Volume I*. London: Pitman.
  - Belkin, N.J. (1980). Anomalous state of knowledge for information retrieval. *Canadian Journal of Information Science*, **5**, 133-143.
  - Belkin, N.J. Cool, C., Stein, A., & Thiel, U. (1995). Cases, scripts and information seeking strategies: on the design of interactive information retrieval systems. *Expert Systems with Application*, **9**(3), 379-395.
  - Bellardo, T. (1985). An investigation of online searcher traits and their relationship to search outcome. *Journal of the American Society for Information Science*, **36**(4), 241-250
  - Borlund, P. (2000). Experimental components for the evaluation of interactive information retrieval systems. *Journal of Documentation*, **56**(1), 71-90.
  - Borlund, P. & Ingwersen, P. (1997). The development of a method for the evaluation of interactive information retrieval systems. *Journal of Documentation*, **53**(3), 225-250.
  - Borlund, P. & Ingwersen, P. (1998). Measures of relative relevance and ranked half-life: performance indicators for interactive IR. In: W.B. Croft, A. Moffat, C.J. van Rijsbergen, R. Wilkinson & J. Zobel (eds.), *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. New York, NY: Association for Computing Machinery: 324-331.
  - Bunge, M.A. (1967). *Scientific research*. (2 vols.) Heidelberg: Springer-Verlag.
  - Byström, K. (1999). *Task complexity, information types and information sources*. Doctoral Dissertation. Tampere: University of Tampere. (Acta Universitatis Tamperensis 688).
  - Byström, K. & Järvelin, K. (1995). Task complexity affects information seeking and use. *Information Processing & Management*, **31**(2), 191 - 213.
  - Campbell, D.J. (1988). Task complexity: a review and analysis. *Academy of Management Review*, **13**(1), 40-52.
  - Daft, R.L. & Sormunen, J. & Parks, D. (1988). Chief executive scanning, environmental characteristics, and company performance: an empirical study. *Strategic Management Journal*, **9**(2), 123-139.
  - Dervin, B. & Nilan, M. (1986). Information needs and uses. *Annual review of information science and technology*, **21**, 3-33.
  - Ellis, D. (1989). A behavioural approach to information retrieval design. *Journal of Documentation*, **46**(3), 318-338.
  - Ellis, D. & Cox, D. & Hall, K. (1993). A comparison of the information seeking patterns of researchers in the physical and social sciences. *Journal of Documentation*, **49**(4), 356-369.
  - Engelbart, D.C. (1962). [Augmenting human intellect: a conceptual framework](#). Menlo Park, CA: Stanford Research Institute. (Summary report AFOSR-3233) Retrieved 27 September 2003 from <http://www.bootstrap.org/augdocs/friedewald030402/augmentinghumanintellect/ahi62index.html>
  - Fabritius, H. (1998). [Information seeking in the newsroom. Application of the cognitive framework for analysis of the work context](#). *Information Research*, **4**(2). Retrieved 27 September, 2003, from <http://informationr.net/ir/4-2/isc/fabritiu.html>
  - Fischer, W.A. (1979). The acquisition of technical information by R&D managers for problem solving in nonroutine contingency situations. *IEEE Transactions on Engineering Management*, **26**(1), 8-14.
  - Fiske, D.W. & Maddi, S.R. (1961). *Functions of varied-experience*. Homewood, IL: Dorsey Press.
  - Greene, F. & Loughbridge, B. (1996). [Investigating the management information needs of academic Heads of Department: a Critical Success Factors approach](#). *Information Research*, **1**(3) Retrieved 27 September, 2003, from <http://informationr.net/ir/1-3/paper8.html>
  - Hackman, J.R. (1969). Toward understanding the role of tasks in behavioral research. *Acta Psychologica*, **31**, 97-128.
  - Hart, P.J. & Rice, R.E. (1991). Using information from external databases: contextual relationships of use, access method, task, database type, organizational differences, and outcomes. *Information Processing & Management*, **27**(5), 461-479.
  - Herner, S., & Herner, M. (1967). Information needs and uses in science and technology. *Annual Review of Information Science and Technology*, **2**, 1-34.
  - Ingwersen, P. (1996). Cognitive perspectives of information retrieval interaction. *Journal of Documentation*, **52**(1), 3-50.
  - Järvelin, K. (1986). On information, information technology and the development of society: an information science perspective. In P. Ingwersen, L. Kajberg, & A. Mark Pejtersen (eds.), *Information technology and information use: towards a unified view of information and information technology*. London: Taylor Graham: 35-55.

- Järvelin, K. (1987). Kaksi yksinkertaista jäsennystä tiedon hankinnan tutkimusta varten [Two simple conceptual frameworks for information seeking research]. *Kirjastotiede ja Informatiikka*, **6**(1), 18-24. [In Finnish, English abstract]
- Järvelin, K. & Repo, A. (1983). On the impacts of modern information technology on information needs and seeking: a framework. In H.J. Dietschmann, (Ed.), *Representation and exchange of knowledge as a basis of information processes* (pp. 207-230). Amsterdam, NL: North-Holland.
  - Järvelin, K. & Repo, A. (1984). A taxonomy of knowledge work support tools. *Proceedings of the Annual Meeting of the American Society for Information Science*, **21**, 59-62.
  - Kernan, J.B., & Mojena, R. (1973). Information utilization and personality. *Journal of Communication*, **23**(3), 315-327
  - Kuhlthau, C.C. (1991). Inside the search process: information seeking from the user's perspective. *Journal of the American Society for Information Science*, **42**(5): 361-371.
  - Kuhlthau, C.C. (1993). *Seeking meaning: a process approach to library and information services*. Norwood, NY: Ablex.
  - McClelland, D.C. (1961). *The achieving society*. New York, NY: Van Nostrand,
  - MacMullin, S.E. & Taylor, R.S. (1984). Problem dimensions and information traits. *The Information Society* **3**(), 91-111.
  - March, J. & Simon, H. (1967). *Organizations*. (2nd ed.). New York, NY: Wiley.
  - Murtonen, K. (1992). Tuloksellisempaan tiedonhankintatutkimukseen: prosessianalyysi tiedontarpeiden ja tiedonhankinnan tutkimuksessa [Toward more effective information seeking studies: use of process-analysis in information needs and information seeking research]. *Kirjastotiede ja Informatiikka* **11**(2), 43-52. (In Finnish)
  - Murtonen, K. (1994). *Ammatilliset tiedontarpeet ja tiedonhankinta tutkimuskohteena: Tutkimus tehtävän kompleksisuuden vaikutuksista tiedontarpeisiin ja tiedonhankintaan*. [Professional information needs and information seeking as study objects: a study on the effects of task complexity on information needs and information seeking]. Thesis for the Degree of Licentiate of Social Sciences. Tampere: University of Tampere, Department of Information Studies. (In Finnish)
  - Palmer, J. (1991). Scientists and information. II. Personal factors in information behaviour. *Journal of Documentation*, **47**(3), 254-275
  - Roberts, K.H. & Glick, W. (1981). The job characteristics approach to task design: a critical review. *Journal of Applied Psychology*, **66**(2), 193-217.
  - Saracevic, T. (1996). Modeling interaction in information retrieval: a review and proposal. *Proceedings of the Annual Academy Meeting of American Society for Information Science*, **33**, 3-9.
  - Siatri, R. (1998). Information seeking in electronic environment: a comparative investigation among computer scientists in British and Greek universities. *Information Research*, **4**(2). Retrieved 27 September, 2003, from <http://informationr.net/ir/4-2/isic/siatri.html>
  - Spink, A. (1997). Study of interactive feedback during mediated information retrieval. *Journal of the American Society for Information Science*, **48**(5), 382-394.
  - Thompson, J.D. (1967). *Organizations in action*. New York, NY: McGraw-Hill Book Co.
  - Tiamiyu, M.A. (1992). The relationships between source use and work complexity, decision-maker discretion and activity duration in Nigerian government ministries. *International Journal of Information Management*, **12**(2), 130-141.
  - Timko, M., & Loynes, R.M.A. (1989). Market information needs for prairie farmers. *Canadian Journal of Agricultural Economics*, **37**, 609-627.
  - Tushman, M.L. (1978). Technical communication in R&D laboratories: the impact of project work characteristics. *Academy of Management Journal*, **21**(4), 624-645.
  - Vakkari, P. (1998). Growth of theories on information seeking. An analysis of growth of a theoretical research program on relation between task complexity and information seeking. *Information Processing & Management*, **34**(3/4), 361-382.
  - Vakkari, P. (1999). Task complexity, problem structure and information actions. Integrating studies on information seeking and retrieval. *Information Processing & Management*, **35**(6), 819-837.
  - Vakkari, P. & Kuokkanen, M. (1997). Theory growth in information science: Applications of the theory of science to a theory of information seeking. *Journal of Documentation*, **53**(5), 497-519.
  - Van de Ven, A. & Ferry, D. (1980). *Measuring and assessing organizations*. New York, NY: Wiley.
  - Wagner, D. & Berger, J. (1985). Do sociological theories grow? *American Journal of Sociologym* **90**, 697-728.
  - Wilson, T.D. (1981). On user studies and information needs. *Journal of Documentation*, **37**(1), 3-15.

- Wilson, T.D. (1997). Information behaviour: an interdisciplinary perspective. *Information Processing & Management*, **33**(4), 551-572.
- Wilson, T.D. (1999). Models in information behaviour research. *Journal of Documentation*. **55**(3), 249-270.
  - Wilson, T.D. & Streatfield, D.R. (1980). [\*"You can observe a lot..." A study of information use in local authority social services departments conducted by Project INISS\*](#) Sheffield: University of Sheffield, Postgraduate School of Librarianship and Information Science. (Occasional Publication No. 12) Retrieved 27 September, 2003, from <http://informationr.net/tdw/publ/INISS/>
  - Wood, R.E. (1986). Task complexity: definition of the construct. *Organizational Behavior and Human Decision Processes* **37**(1), 60-82.
  - Zeffane, R.M. & Gul, F.A. (1993). The effects of task characteristics and sub-unit structure on dimensions of information processing. *Information Processing & Management*. **29**(6): 703-719.
- 

## Appendix

Byström's (1999) eleven statements - cf. Figure 5.

S1: as soon as information acquisition requires an effort people as sources are more popular than documentary sources.

S2: the more information types are needed, the greater the share of people as sources.

S3: the more information types are needed, the greater the share of general-purpose sources and the smaller the share of task-oriented sources.

S4: the more information types are needed, the more sources are used.

S5: the internality of different source types is loosely connected to the information types.

S6: the higher the degree of task complexity, the more probable is the need for multiple information types: first task information, then task and domain information, and finally task, domain and [problem] solving information.

S7: the higher the degree of task complexity, the more information types are needed, and the greater the share of people as sources and the smaller the share of documentary sources.

S8: the higher the degree of task complexity, the more information types are needed and the greater the share of general-purpose sources and the smaller the share of task-oriented sources.

S9: the higher the degree of task complexity, the more information types are needed, and the higher the number of sources used.

S10: task complexity is distinctly related to increasing internality of people as sources and decreasing internality of documentary sources.

S11: Increasing task complexity fosters the use of people as sources.

---

Find other papers on this subject.

---

Check for citations, [using Google Scholar](#)

---

**54338**  
[Web Counter](#)

© the authors, 2003.  
Last updated: 29 September, 2003

- 
- [Contents](#) |
  - [Author index](#) |
  - [Subject index](#) |
  - [Search](#) |
  - [Home](#)
-