

This is a preprint of an article accepted for publication in Journal of the Association for Information Science and Technology copyright © 2016 (Association for Information Science and Technology). This reprint has been updated to reflect changes in the final version.

# Search task features in work tasks of varying types and complexity

**Miamaria Saastamoinen and Kalervo Järvelin**

University of Tampere, School of Information Sciences, FIN-33014 Tampere, Finland. E-mail: {miamaria.saastamoinen, kalervo.jarvelin}@staff.uta.fi

## Abstract

Information searching in practice seldom is an end in itself. In work, work task (WT) performance forms the context, which information searching should serve. Therefore information retrieval (IR) systems development/evaluation should take the WT context into account. The present paper analyzes how WT features: task complexity and task types, affect information searching in authentic work: the types of information needs, search processes and search media. We collected data on 22 information professionals in authentic work situations in three organization types: city administration, universities and companies. The data comprise 286 WTs and 420 search tasks (STs). The data include transaction logs, video recordings, daily questionnaires, interviews and observation. The data were analyzed quantitatively. Even if the participants used a range of search media, most STs were simple throughout the data, and up to 42 % of WTs did not include searching. WT's effects on STs are not straightforward: different WT types react differently to WT complexity. Due to the simplicity of authentic searching, the WT/ST types in interactive IR experiments should be reconsidered.

## Introduction

Much of modern work is information-intensive and supported by diverse information systems. Information supply and searching are key activities in information-intensive work. Work tasks (WTs) are the building blocks of work. A job description includes abstract WTs that manifest as WT actions in daily work (Byström & Hansen, 2005). In working life, WT performance forms the context, which information supply and searching should serve. In order to design information search systems to properly serve WT requirements, it is necessary to study how the WTs as actions are connected to searching. Therefore information (retrieval) systems development and evaluation should not take place in isolation but take the work context into account and find out for what purposes and how the systems are used. Failing to do this may result in developing suboptimal systems for expected but biased search needs. In the present paper we aim to find out, how searching is carried out in work contexts and what affects it.

Several researchers highlight the importance of task-based information searching (TBIS) as a research area (Järvelin et al., 2015; Kelly, Arguello, & Capra, 2013; Vakkari, 2003). However, naturalistic field studies of TBIS are rare. Other types of empirical approaches to TBIS seem more common, including log analyses (e.g., Kotov, Bennett, White, Dumais, & Teevan, 2011), simulated WTs (e.g., Borlund, 2003), self-report methods (e.g., Li, 2009), and semi-naturalistic studies (e.g., Pharo, 2004).

The present study belongs to the naturalistic field studies, where information searching is studied in the context of authentic WTs, at the work place, collecting multiple types of data in real-time. There are some predecessors to the present study, e.g., Hansen's (2011) study on TBIS in the patent domain; Saastamoinen, Kumpulainen, and Järvelin's (2012) study in the public administration domain; and Kumpulainen's (2014) study in Molecular Medicine. Also Kellar, Watters, and Shepherd (2006) and Kelly (2006) used naturalistic methods but their work was more concerned with leisure time activities than WTs.

We consider major search task (ST) features that have received little attention in past research, under the influence of WTs. First, information searching is typically studied in a limited context, where the searcher has only *one given search tool* to use. In reality, there are several options the searcher can choose from. Second, in experiments, the searcher cannot choose to use *more than one search tool* though this is often possible in real-life. Third, *information needs* are typically studied narrowly: Either the need is almost directly given to searchers (in the case of simulated WTs), or the searchers are asked to describe their needs (in self-report studies), or researchers attempt to reconstruct information

needs based on a sequence of queries and clicks (in log analyses). These approaches are unable to show the existence and/or prevalence of these information needs in reality. Fourth, modeling *search processes* is often neglected in favor of result list ranking and relevance of individual documents. Fifth, *participants* are often students, or employees of a single organization. Thus the generalizability of the findings concerning their practices may be limited.

To overcome these limitations, we propose analyzing information need formation and search tool use in naturalistic STs in authentic and heterogeneous WT context. This will shed light on whether current TBIS research addresses issues prevalent in authentic STs/WTs. Furthermore, real ST processes should be analyzed if we want to understand and support them better. We think that it is necessary in information searching experiments to study heterogeneous search environments and WTs; or, the WTs have to be problem-centered rather than search-centered.

The present paper studies perceived WTs and their subtasks, STs, that were inferred from the data. Each ST has one underlying information need, a driving force to begin searching. The independent variables are *WT type*, a cross-domain data-driven categorization, and (*perceived*) *WT complexity*, a traditional and significant variable in earlier TBIS studies (e.g., Byström & Järvelin, 1995). The dependent variables are the ST features: information need and ST process, the main search medium, and the number of unique search media. The present study is unique in its *combination* of the amount of authentic data collected and its diversity, the quantitative approach and the dependent variables.

Altogether 22 participants in three types of organizations were followed performing 286 WTs including 420 STs. The data include transaction logs, questionnaire responses, observation, screen video recordings and interviews. These features allow the study of issues often neglected in earlier studies: what search media are used if people can freely choose among those available and whether there are several of them; what are the participants' authentic information needs and how the participants act to fulfill them.

The rich data set also includes other analyzable search features, e.g., concerning the queries. For reasons of readability and length, findings concerning these features are discussed in another paper (Saastamoinen & Järvelin, 2016).

## Related research

### *Concepts*

*Work tasks* (WTs) are formed of the actions performed in order to achieve a goal, the task outcome (Vakkari, 2003). *Search tasks* (STs) are subtasks of WTs. Various WT and ST features are discussed in information studies literature (Li & Belkin, 2008).

WTs are often classified in TBIS studies based on their *complexity* and *difficulty*. These two concepts are neither identical nor independent of each other. (Liu, Kim, & Creel 2015; Wildemuth, Freund, & Toms, 2014). Another classification divides complexity into *objective* vs. *subjective* measures (Liu & Li, 2012). Liu and Li (2012) argue that these terms are often confused while Gwizdka and Spence (2006) and Maynard and Hakel (1997) found that they nevertheless correlate. In the present study, we apply subjective WT complexity. Without in-depth substance knowledge of the study participants' tasks, assessing objective complexity would have been prone to mistakes if not impossible.

*Information needs* induce STs. We apply an information need categorization similar to Broder's (2002) but use it to describe the intent behind the whole ST instead of a single query. Broder's (2002) taxonomy includes navigational, informational and transactional searches. We classify information needs into instrumental (the intention is to perform an action or navigation) and content-driven (new information needed); and in the latter case, into factual, topical, or known item needs (cf. Ingwersen, 1986; Ingwersen & Järvelin, 2005; Toms, 2011).

Kumpulainen's (2014) search trails inspired our classification of *ST performance processes*. She analyzed five types of trails. Our classification is similar except for one type, 'chain', which seemed specific to the field she studied, molecular medicine. Our classes were identified in the data and are: single, list, stable and developing process. These are explained in detail in the Analysis section.

### *Approaches to TBIS*

There are several empirical approaches to TBIS. Next, they are briefly discussed.

Log analyses exploiting search engine log data are unbeatable in the quantity of data. However, logs do not offer factual knowledge about the underlying WTs. One may try to reconstruct the STs based on the log (e.g., Kotov et al., 2011) but without further information these are mere approximations. Furthermore, typical log studies include only queries to a

single system, while different (types of) systems may be used in authentic WTs. However, if a log study is conducted in a naturalistic environment with background information, it may provide a great contribution for TBIS. For example, Kellar et al. (2006) collected logs of students' Web usage in their own computers and found four main ST types that affected search behavior.

Simulated WTs (Borlund, 2003) are pre-constructed STs that have a short background story providing motivation and a frame for searching. They work as simulations of real situations when they are carefully designed, so that the participants can identify with the task at hand. However, the veracity of simulated WTs is left to the imagination of the researcher.

Self-report methods employ diaries (Du, 2014), surveys (Nicholas, Williams, Rowlands, & Jamali, 2010) or interviews (Li, 2009) in data collection. Self-report methods are quite straightforward to apply but their results are not always reliable because participants may report their actions inaccurately.

In semi-naturalistic studies, information needs and tasks may be authentic but information searching occurs in a controlled environment (e.g., Pharo & Järvelin, 2004; Vakkari, Pennanen, & Serola, 2003). This supports control of data collection but significantly limits the participants' choices in TBIS.

Naturalistic field studies are rare in information searching research, perhaps because they require plenty of resources. In only a few studies (e.g. Hansen, 2011; Kumpulainen, 2014; Saastamoinen et al., 2012) researchers arrived at the work places of participants to follow their authentic work rather than bringing the participants into a controlled environment to perform assigned tasks. However, often the number of participants in the studies is low and little if any quantitative results are gained. The main findings of naturalistic studies of TBIS include:

- When WTs become more complex, more information system integration is needed, and search situations become more complex because of lack of automated integration and the need to generate manual workarounds (Kumpulainen & Järvelin, 2010; Kumpulainen, 2014).
- Information needs often develop during an ST (Hansen, 2011). This confirms earlier views by Bates (1989) and Borlund (2003).
- Along increasing WT complexity, the use of network sources (including the Web) increases, whereas the use of organizational information systems decreases (Saastamoinen et al., 2012).

Past studies indicate that tasks affect the use of information systems and need for their integration. Multiple information systems are used in WTs. These are findings which are hardly taken into account in traditional information searching studies but which are clearly visible in the field. At best, these approaches complement each other and any single study type is insufficient alone.

## **Research design**

### *Research questions*

Earlier studies have focused on either maximizing quantitative data, but largely ignored context, or on maximizing qualitative data by analyzing the actions of a few participants in great detail, but excluded the possibilities for generalization. Moreover, asking participants what they did does not yield quite as reliable data as one may collect by observing and logging their actions. We aim at a large, quantitative analysis of TBIS with focus on WT factors affecting search. This helps design better information systems and better experimental studies in the future.

On the basis described above, we study four STs features: the main search medium, the number of unique search media, the types of information needs and ST processes in each of the following research questions:

- 1 How does subjective WT complexity affect STs?
- 2 How do WT types affect STs?
- 3 What is the combined effect of WT type and complexity on STs?

### *Data collection*

The data collection took place in authentic working environments. All WTs are authentic. Twenty-two knowledge workers participated in the study. They worked in six organizations; a city administration (10 participants), two commercial companies (7) and three universities (5). The participants were recruited through personal connections to these organizations. The participants got two movie tickets per person (for those who had participated for several days). Figure 1 presents the data collection protocol.

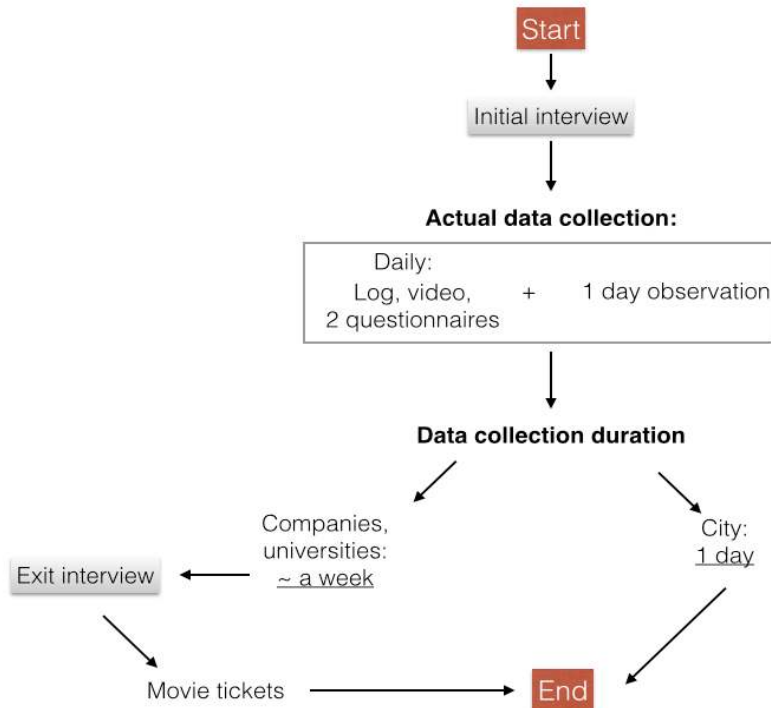


FIGURE 1. The data collection protocol.

The data were collected using multiple methods. The first step was a free-form recorded opening interview with each participant, aiming to give the researchers an understanding of the participants' work and help create a confidential relationship between the parties. The participants were asked to tell about their WTs, information sources, and search practices in particular. After the interview, we agreed on the days of actual data collection. The settings of the data collection software were also checked and instructions given on their use. The study software was installed by the technical support of each organization.

The duration of data collection varied by participant and organization from one to 11 days. On each day, participants filled in an electronic questionnaire (Appendices 1-2) when beginning to work and just before leaving from work. In the morning, they filled in the expected WTs of the day, their complexity, and the information systems and other sources they expected to use. In the afternoon, they listed the WTs performed, their complexity and the information sources used.

Information interaction during the working days was recorded with a logging software (Figure 2) and a screen capture software. The software did not cause serious problems in computer performance. The participants were given instructions about the data collection procedure in writing and orally with continuous support offered by email/phone. To preserve confidentiality, the participants were able to stop automatic logging, delete rows from the log, or put 'private' tags on the log for the researcher to delete confidential data. Some participants used these options. After removal of the confidential sections, the logs consisted of 40 200 rows.

| Name  | Start          | End            | Duration | Process               |
|---|----------------|----------------|----------|-----------------------|
| Microsoft Excel - Selection 1-5 2014 READY.xls [Compatibility mod | 29.01.14 14:23 | 29.01.14 14:24 | 00:00:06 | Microsoft Office 2010 |
| GLINK   | 29.01.14 14:24 | 29.01.14 14:24 | 00:00:05 | gl                    |
| Microsoft Excel - Selection 1-5 2014 READY.xls [Compatibility mod | 29.01.14 14:24 | 29.01.14 14:24 | 00:00:06 | Microsoft Office 2010 |
| GLINK   | 29.01.14 14:24 | 29.01.14 14:27 | 00:03:16 | gl                    |
| No topic - Message (HTML)   | 29.01.14 14:27 | 29.01.14 14:27 | 00:00:18 | Microsoft Outlook     |

FIGURE 2. An example of the log exported from the logging software. All original Finnish texts are translated in the

figure for clarity reasons. Any confidential information has been removed. 'Glink' refers to an emulator whose name appears when the participant is using an organizational information system.

One data collection day involved a free-form direct observation, lasting for the whole working day. The participant was advised to work normally. The researcher sat next to her, observing her work, using voice-recording and pen-and-paper to record the participant's actions. The researcher asked questions for clarification, such as, 'Which WT are you doing now?', and 'What are you searching for?'. Notes were taken when necessary, especially on non-computer-based actions. Observation supported greatly understanding the events in the log and, therefore, the reliability of the analysis. A multi-day observation would certainly have provided richer data but this would have severely complicated the recruitment process.

After the data collection finished, the researcher met the participant to fetch the log data. The participant asked the technical support of her organization to uninstall the software used.

An exit interview was held a few weeks after the data collection phase with the 12 participants, whose data collection lasted more than one day. Before the interview, the researcher performed a preliminary data analysis. This consisted of identifying WTs and interesting or unclear queries, which were then discussed in the interview. Typical questions were whether a WT mentioned in the morning questionnaire but missing in the afternoon was deliberately omitted or just overlooked; whether a row in the log was part of some WT; where the boundaries between WTs were; and what was the reason for a query. The participants were shown parts of the log or the screen video when necessary. The interview was recorded.

Table 1 shows the basic distributions of the data between organization types. We were allowed to collect data only one day per participant in the municipal administration. However, this was not considered as an issue since we do not analyze differences between organizations. Rather, each participant represents her own WTs.

TABLE 1. Data overview. Minimum and maximum by participant in parentheses (min-max).

|                      | <b>Companies</b>                       | <b>City</b>                                    | <b>Universities</b> |
|----------------------|--|--|---------------------|
| Participants         | 7                                      | 10   | 5                   |
| Sample job roles     | Design manager, marketing professional | Communications manager, communications officer | Teacher, researcher |
| Data collection days | 37 (5-6)                               | 10 (1-1)                                       | 30 (4-11)           |
| WTs                  | 138 (14-28)                            | 47 (3-6)                                       | 101 (14-36)         |
| STs                  | 192 (13-48)                            | 55 (1-11)                                      | 173 (6-97)          |

### *Analysis*

For the analysis, we had electronic questionnaire responses, logs, screen capture, observation notes and interview data to answer our research questions. The preliminary analysis conducted before the exit interviews provided a useful basis for the analysis. Below we describe initial WT identification, WT type and complexity categorizations, ST identification and feature categorization, and statistics. Table 2 provides an overview on how the data were used.

TABLE 2. Data types.

| <b>Data type</b>      | <b>Function</b>  |
|-----------------------|--|
| Initial interview     | <ul style="list-style-type: none"> <li>- understanding the work of the participant</li> <li>- telling the participant about the study</li> <li>- agreeing on details about data collection</li> </ul>  |
| Morning questionnaire | <ul style="list-style-type: none"> <li>- two out of three WT complexity estimates</li> <li>- task descriptions as a basis for WT type classification</li> <li>- information about which tasks are likely to be performed</li> <li>- potential information source use; helps to spot the WTs in the logs</li> </ul> |
| Transaction log       | <ul style="list-style-type: none"> <li>- concrete steps of work, divided into WTs and STs by the researcher</li> <li>- the number of unique search media in STs</li> <li>- basis for the classification of ST features</li> </ul>  |
| Screen video          | <ul style="list-style-type: none"> <li>- concrete steps of work in visual form</li> <li>- understanding the textual log</li> </ul>   |

|                         |  |
|-------------------------|--|
|                         | <ul style="list-style-type: none"> <li>- spotting WTs and STs</li> <li>- support for the selection of suitable search feature categories for each ST</li> </ul>  |
| Observation             | <ul style="list-style-type: none"> <li>- support for interpreting the log</li> <li>- reliably spotted WTs</li> <li>- understanding the ways participants work</li> </ul>   |
| Afternoon questionnaire | <ul style="list-style-type: none"> <li>- one out of three WT complexity estimates</li> <li>- task description as a basis for WT type classification</li> <li>- information about which tasks were performed</li> <li>- information source use and estimated time of performance helping to spot WTs in the logs</li> </ul> |
| Exit interview          | - knowledge about WT boundaries, STs and unclear parts in the log  |

---

We started by identifying the WTs in the data. The morning and afternoon questionnaires of each day were compared to each other in order to link the WTs that were mentioned in both, and find which ones were mentioned only in one of them. The next step was to find these tasks in the log and screen video. The clues were collected from all data types, and the exit interview helped a lot. Some tasks could not be found because they were not actually performed, others because they were performed but did not include any use of a computer (e.g., some meetings) and were thus excluded from the analysis. Otherwise, all the WTs mentioned in the questionnaire(s) and recognizable in the log or in the video were included in the analysis.

We did not need to give the participants detailed instructions about how to describe or define their WTs since they were capable of doing that in the first interview. Indeed, the questionnaire responses show that people across different fields and work roles were able to similarly recognize the WTs that they believed forming their work day, and name them recognizably and at the same level of abstraction.

#### *WT complexity and type*

Our WT complexity measure is formed of three figures, each on a scale from 0 to 100 percent, as estimated by the participants: a) pre-task complexity; b) post-task complexity; c) amount of prior knowledge of task process. The complexity is the average of a, b, and the complement of c calculated using formula  $(a+b+(100-c))/3$ . As an example, if a participant estimates a WT to have 20% complexity in the morning (a), 10% in the afternoon (b), and she estimates her prior knowledge of task process to be 90% (c), the WT complexity is calculated as  $(20+10+(100-90))/3 = 13.33\%$ . In case of any missing estimates, the above formula was used as an average of the remaining components. Similar measures have been successfully used earlier (e.g., Byström & Järvelin, 1995; Kumpulainen, 2014; Saastamoinen et al., 2012). Using the average of three estimates better represents the underlying WT complexity. For example post-task complexity is prone to the effects of any problems occurring at the end of the WT.

In addition to the continuous WT complexity variable, we categorized WT complexity to allow more analysis types. We categorized WTs into four (as close as possible) equally large categories. The categories are I (tasks from 0% complexity to 21.7%), II (21.8-38.3%), III (38.4-50%) and IV (50.1-100%). This categorization allows us to study the effects of relative complexity in the data.

After data collection, we saw that also another dimension of WTs would help elaborate the effects of WT complexity. However, earlier studies did not provide us with a ready suitable categorization. By carefully reading the task descriptions, we created a tailor-made categorization. The classification reliably separates task types from each other, potentially relates to information searching and use, is not inherently related to task complexity, suits all organizations studied, and divides the data into categories of roughly the same size in order to enable comparisons between the task type groups. The process of forming these two categorizations, WT type and complexity, were independent of each other: neither categorization affected the other one. The types are as follows.

In *communication tasks*, the communication scope stands out in the description: it includes informing, being informed or both, e.g., 'Handling emails'. This category is similar to collaboration or group tasks studied by Foster (2006).

*Support tasks* are often administrative but they can be other (almost) mechanical WTs as well. WTs that include several mixed smaller tasks are also considered as support tasks. Support tasks may coincide with routine tasks, a task type presented by for example, Li and Belkin (2008) but our support tasks do not need to be frequent.

*Editing and intellectual tasks* include a creative aspect. The difference is that in intellectual WTs, participants create something from the beginning, whereas in editing WTs, the groundwork is already done, or the objective is to make

only final edits. Intellectual tasks typically require cogitation as in the case of writing a report or making a significant decision. Editing tasks include finishing a text, commenting on texts or giving marks to students. Similar task types to our editing and intellectual tasks are Algon's (1997) 'report generation', Hackman's (1968) 'intellective tasks', and Li and Belkin's (2008) 'intellectual tasks'. Our support, editing and intellectual tasks form a continuum similar to task complexity, later called the *sup-edit-int continuum*.

Intra- and inter-classifier reliability tests concerning WT type classification were performed. The average agreement of two inter-classifier tests was 78% (Cohen's Kappa .69). Intra-classifier reliability was 88% (Cohen's Kappa .83).

### *ST features*

We defined STs as subtasks fulfilling two necessary conditions: an ST has a somewhat unified 'information need', i.e. an underlying reason to search (a goal), and contains at least one query. STs cannot overlap. Mere browsing a directory (without a query) does not comprise an ST. The identification of STs within WTs was far from trivial since STs are self-initiated and self-constructed by the searchers. The STs do not manifest themselves through explicit labels like in the logs of interactive information retrieval (IIR) experiments.

A query is identifiable either in the log or in the screen capture video or both. A query contains characters written by the participant and a command to search. Queries can also be issued by selecting pre-defined values in a drop-down menu or by checking menu boxes and giving a search command.

Each ST includes querying in some *search media*. Search medium as concept is similar to the information source discussed in several papers. The difference is that search media are sources used by querying, not for example by browsing. We identified four categories of search media: communication channels (e.g. email), organizational databases, the Web, and personal computer (PC) (e.g. directory search). We also identified unique search media within STs. For example, a search engine and a Web dictionary are two unique media while belonging to the same main search media category, the Web. The main search medium is the one that received the greatest number of queries within an ST.

Our approach to *information needs* here is quite practical: they represent the goal of the ST rather than the mental stage of the searcher. It is possible that the mental stage of the searcher develops during the ST, i.e. the intent behind each query may be a little different (see also ST performance processes below). However, our information need categorization is meant to describe the intent behind the whole ST, not individual queries. We identified the four information need categories: factual, such as a name; a known item when looking for a unit of information, such as a file, known or expected to exist; topical when the searcher wants to know more than a fact; or instrumental, when the searcher wants to accomplish something, such as navigation.

Each ST is realized as a process, the *ST performance process*. STs with one query belong to the category *single process*. A *list process* means a sequence of independent queries, where each query has its own but related information need motivating it, such as checking whether the telephone numbers in a file are correct. In a *stable process*, the intent behind the search is a coherent whole. The results collected during the process do not essentially affect the queries. The list and stable processes differ in that the connection between queries in a list is more of a technical matter while in stable processes the connection lies more in the contents. In a stable process, the searcher may have a topic in mind that she is able to find in smaller pieces for aggregation. In a *developing process*, queries build on each other and the information in the previous results.

For further analysis, we organized information needs from the most simple to the most complex, that is, instrumental, factual, known item, and topical. ST performance processes were similarly ordered, that is single, list, stable and developing. These are called information need complexity and performance process complexity classifications.

All ST features presented above are close to the ones common in the literature. We have adapted them slightly to our data. This was necessary in order to describe information searching in authentic WTs. We sought to keep the modifications minor to serve the comparability of our findings with prior studies.

### *Statistical analysis*

Simple statistical methods were used in the analysis. Though our data set was a convenience sample, the tasks studied can be considered as representing at least some features common to tasks beyond the data. After data collection, WTs and ST features were classified making the data well-structured. The rather large size of the well-structured data allows the use of statistics to support the interpretation of the findings. We use Pearson correlation ('r') for continuous variables, Spearman correlation (' $\rho$ ') if a variable is ordinal, and we perform chi-squared tests (' $\chi^2$ ') for cross tabulations. Statistical significance ('p') is tested for all analyses; we set the limit for significance at  $p=.05$ .

Because STs or WTs were not controlled in any way, the data are authentic. This provides 'noisy' data, i.e., direct, measurable connections between work and ST features may be weak due to intervening uncontrolled factors. However, the connections found, or their absence, are important because they exist in authentic environments. The connections can be studied further in controlled experiments, which is beyond the scope of the present paper. For reasons of clarity and brevity, only the results of the most interesting statistical analyses are reported alongside the findings.

## Results

We first discuss the overall findings, then the effects of WT complexity on ST features followed by a similar section, on results concerning WT types' effects on STs. Finally, the combined effects of WT types and complexity on searching are analyzed.

### Data profile

The data consist of 286 WTs. Their complexity varied between 0% and 90% (mean 39.5%, *SD* 20.3). Individual *per participant* task complexity averages varied between 3.3% and 56.4% (minimums 0%-47.3%, maximums 6.7%-90%, *SDs* 3.4-28). Communication tasks are the most common, and support tasks the least common WT type. Table 3 shows the distribution of WT types across WT complexity categories. There seems to be a connection between WT type and complexity.

TABLE 3. WT types and WT complexity (%).

| Work task type | Work task complexity |                  |                 |                 | total | N   |
|----------------|----------------------|------------------|-----------------|-----------------|-------|-----|
|                | I<br>0-21.7%         | II<br>21.8-38.3% | III<br>38.4-50% | IV<br>50.1-100% |       |     |
| communication  | 20.7                 | 28.3             | 35.9            | 15.2            | 100.0 | 92  |
| support        | 34.1                 | 22.7             | 27.3            | 15.9            | 100.0 | 44  |
| editing        | 26.1                 | 29.0             | 23.2            | 21.7            | 100.0 | 69  |
| intellectual   | 13.6                 | 13.6             | 28.4            | 44.4            | 100.0 | 81  |
| <b>total</b>   | 22.0                 | 23.4             | 29.4            | 25.2            | 100.0 |     |
| <b>N</b>       | 63                   | 67               | 84              | 72              |       | 286 |

The data includes 420 STs. Table 4 shows the distributions of ST feature categories. The *main search media* are quite evenly distributed between organizational databases, the Web and communication media. About 3% of STs occurred mainly on the task performer's PC. Only one search medium is used in most STs (86%) and 99.5% of STs have 1-3 search media, overall average being 1.18. Topical *information needs* form under 20% of STs, factual needs almost a third, and other needs between 20%-30%. Single *search processes* are in the majority (51%), 14% of ST processes are stable, 17% developing, and 18% lists. Moreover, *information need complexity* and *search process complexity* are weakly while significantly correlated ( $\rho=.13, p=.008$ ): the more complex the need, the more complex the process.

TABLE 4. Distributions of ST features ( $N=420$ ).

| ST feature                 | category             | % of STs |
|----------------------------|----------------------|----------|
| <b>main search medium</b>  | <i>pc</i>            | 2.9      |
|                            | <i>org_database</i>  | 33.1     |
|                            | <i>Web</i>           | 33.8     |
|                            | <i>communication</i> | 30.2     |
|                            | <b>total</b>         | 100.0    |
| <b>unique search media</b> | <i>one</i>           | 85.7     |
|                            | <i>more</i>          | 14.3     |
|                            | <b>total</b>         | 100.0    |
| <b>information need</b>    | <i>instrumental</i>  | 21.2     |
|                            | <i>fact</i>          | 32.6     |



|                            |                   |       |
|----------------------------|-------------------|-------|
|                            | <i>known item</i> | 28.1  |
|                            | <i>topic</i>      | 18.1  |
|                            | <b>total</b>      | 100.0 |
| <b>performance process</b> | <i>single</i>     | 51.0  |
|                            | <i>list</i>       | 17.9  |
|                            | <i>stable</i>     | 13.8  |
|                            | <i>developing</i> | 17.4  |
|                            | <b>total</b>      | 100.1 |

### *STs and WT complexity*

The main search media seem independent of WT complexity ( $\chi^2=4.1, p=.902$ ) (Table 5). Organizational databases, the Web and communication media share quite evenly the role of main search medium regardless of WT complexity. The PC always has a share of 2%-3%. Since the independence was surprising, we organized the search media according to their potential flexibility: PC is quite limited in scope since it includes only local programs and files, organizational databases were next, then the Web and last communication media since people as sources of information tend to adapt well to varying information needs. An interesting, small correlation was found between search media and WT complexity. The more complex the WT, the less adaptable media are used ( $\rho=-.10, p=.041$ ). Communication media reach their smallest share (26%) in the most complex WTs. In the same tasks, organizational databases have the largest share, 36%.

TABLE 5. Main search media (%) in varying WT complexity levels.

|                             |       | <b>main search medium</b> |                     |             |                      |              |                       |
|-----------------------------|-------|---------------------------|---------------------|-------------|----------------------|--------------|-----------------------|
|                             |       | <u>PC</u>                 | <u>org_database</u> | <u>Web</u>  | <u>communication</u> | <u>total</u> | <u>n search tasks</u> |
| <b>work task complexity</b> | I     | 1.7                       | 33.9                | <b>35.6</b> | 28.8                 | 100.0        | 59                    |
|                             | II    | 2.3                       | 28.7                | <b>37.9</b> | 31.0                 | 99.9         | 87                    |
|                             | III   | 3.4                       | <b>32.7</b>         | 29.9        | 34.0                 | 100.0        | 147                   |
|                             | IV    | 3.1                       | <b>36.2</b>         | 34.6        | 26.0                 | 99.9         | 127                   |
|                             | total | 2.9                       | 33.1                | <b>33.8</b> | 30.2                 |              | N 420                 |

The number of unique search media in an ST is skewed (with 86% of STs having only one medium). This shows that STs are typically well focused. However, the most complex WTs differ from others. Among them, 21% of STs have more than one search medium while in simpler WTs this varies between 10-14%. Figure 3 shows the average number of unique media across WT complexity.

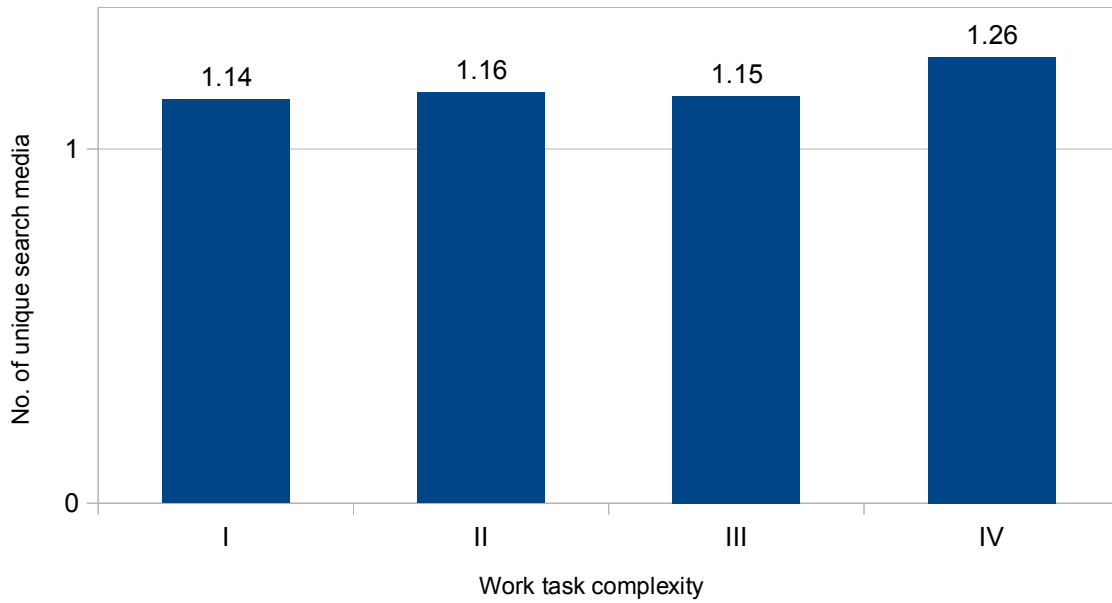


FIGURE 3. The average number of unique search media by WT complexity. Note that the average cannot be  $<1$ , since an ST always includes at least one search medium.

The share of topical information needs is at its largest (24%) in the most complex WTs. The share of instrumental needs is at its largest (32%) in the most simple WTs (Table 6). Interestingly, the share of factual needs is similar in the most simple and the most complex WTs. The shares of known item needs seem independent of WT complexity. Information need complexity is independent of WT complexity ( $q=.07$ ,  $p=.16$ ).

TABLE 6. Information needs (%) in WT complexity levels.

|                      |       | information need |             |             |       |       |                       |
|----------------------|-------|------------------|-------------|-------------|-------|-------|-----------------------|
|                      |       | instrumental     | fact        | known item  | topic | total | <i>n</i> search tasks |
| work task complexity | I     | <b>32.2</b>      | 23.7        | 28.8        | 15.3  | 100.1 | 59                    |
|                      | II    | 23.0             | <b>40.2</b> | 21.8        | 14.9  | 99.9  | 87                    |
|                      | III   | 15.6             | <b>38.8</b> | 29.9        | 15.6  | 99.9  | 147                   |
|                      | IV    | 21.3             | 24.4        | <b>29.9</b> | 24.4  | 100.0 | 127                   |
|                      | total | 21.2             | <b>32.6</b> | 28.1        | 18.1  |       | <i>N</i> 420          |

Single performance processes dominate the data (Table 7); single processes have the largest share in all WT complexity classes. Thus any trends are hard to find between the four task complexity categories and ST performance processes ( $\chi^2=14.1$ ,  $p=.120$ ). However, the share of developing processes reaches its peak (24%) in the most complex WTs and the single processes reach their minimum share in the most complex WTs. The share of list processes grows a little with task complexity. Performance process complexity is connected to WT complexity ( $q=.10$ ,  $p=.037$ ).

TABLE 7. Search process types (%) in WT complexity levels.

|                             |       | search task performance process |             |               |                   |              |                       |
|-----------------------------|-------|---------------------------------|-------------|---------------|-------------------|--------------|-----------------------|
|                             |       | <u>single</u>                   | <u>list</u> | <u>stable</u> | <u>developing</u> | <u>total</u> | <u>n search tasks</u> |
| <b>work task complexity</b> | I     | <b>47.5</b>                     | 15.3        | 16.9          | 20.3              | 100.0        | 59                    |
|                             | II    | <b>60.9</b>                     | 16.1        | 14.9          | 8.0               | 99.9         | 87                    |
|                             | III   | <b>54.4</b>                     | 18.4        | 10.9          | 16.3              | 100.0        | 147                   |
|                             | IV    | <b>41.7</b>                     | 19.7        | 15.0          | 23.6              | 100.0        | 127                   |
|                             | total | <b>51.0</b>                     | 17.9        | 13.8          | 17.4              |              | N 420                 |

*STs in varying WT types*

WT type seems to be connected to the main search medium ( $\chi^2=30.7, p<.001$ ) (Table 8). PC is a rare main search medium but it is most used for searching in intellectual WTs. The use of organizational databases stays quite steady over WT types. The Web reaches its largest share in support tasks (44%), whereas it is seldom used as the main search medium in communication tasks. Communication media are used often as the main medium when searching for information in communication tasks (44%). Communication media are used the least as main media for searching in intellectual WTs (19%).

TABLE 8. Main search media (%) in varying WT types.

|                       |                      | main search medium |                     |             |                      |              |                       |
|-----------------------|----------------------|--------------------|---------------------|-------------|----------------------|--------------|-----------------------|
|                       |                      | <u>PC</u>          | <u>org_database</u> | <u>Web</u>  | <u>communication</u> | <u>total</u> | <u>n search tasks</u> |
| <b>work task type</b> | <u>communication</u> | 0.8                | 34.8                | 20.5        | <b>43.9</b>          | 100.0        | 132                   |
|                       | <u>support</u>       | 3.6                | 27.3                | <b>43.6</b> | 25.5                 | 100.0        | 55                    |
|                       | <u>editing</u>       | 2.0                | 33.3                | <b>35.4</b> | 29.3                 | 100.0        | 99                    |
|                       | <u>intellectual</u>  | 5.2                | 33.6                | <b>41.8</b> | 19.4                 | 100.0        | 134                   |
|                       | <u>total</u>         | 2.9                | 33.1                | <b>33.8</b> | 30.2                 |              | N 420                 |

STs in all WT types are mostly conducted with one unique search medium. The sum of unique search media over the 420 STs is only 497. In communication tasks, even 95% of STs have only one unique search medium which indicates that searching in communication WTs is especially well focused. Intellectual WTs have the largest number of unique media, more than one search medium in 25% of STs. Figure 4 shows the average number of unique search media across WT types.

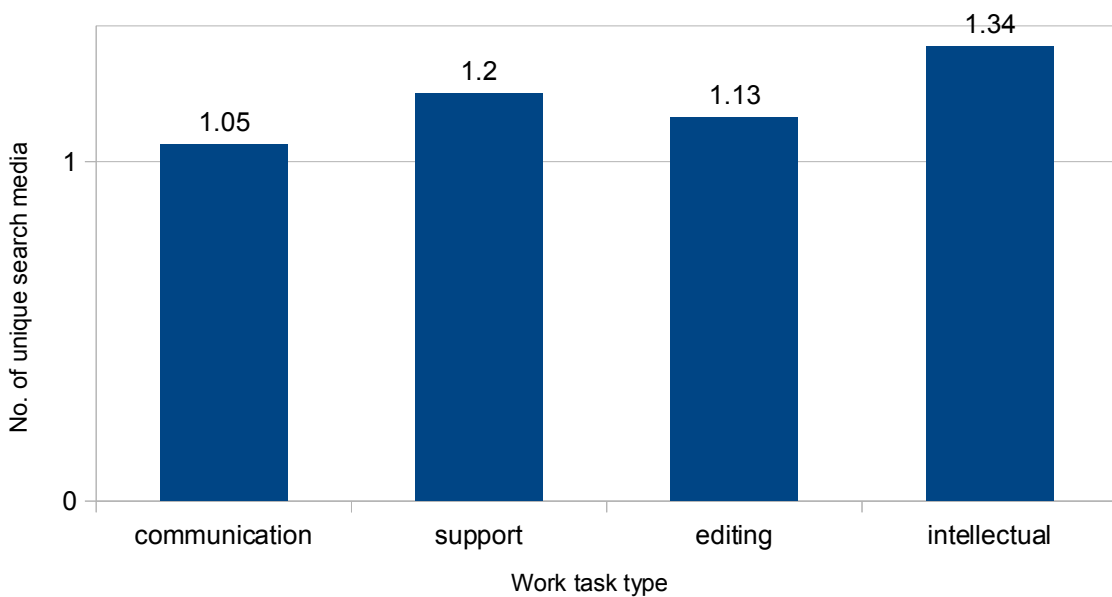


FIGURE 4. The average number of unique search media across WT types. Note that the average cannot be <1, since an ST always includes at least one search medium.

WT type is connected to information need ( $\chi^2=37.3$ ,  $p=.001$ , Table 9). In communication WTs, instrumental needs are most common and more common than in other WT types, and topical searches are rare. Factual needs are less frequent than in other WT types. Along the WT continuum *sup-edit-int*, the shares of topical and known item needs increase and the shares of factual and instrumental needs decrease.

TABLE 9. Information needs (%) in WT types.

|                |               | information need |             |             |       |       | total | n search tasks |
|----------------|---------------|------------------|-------------|-------------|-------|-------|-------|----------------|
|                |               | instrumental     | fact        | known item  | topic |       |       |                |
| work task type | communication | <b>31.8</b>      | 27.3        | 25.8        | 15.2  | 100.1 | 132   |                |
|                | support       | 29.1             | <b>40.0</b> | 21.8        | 9.1   | 100.0 | 55    |                |
|                | editing       | 22.2             | <b>36.4</b> | 27.3        | 14.1  | 100.0 | 99    |                |
|                | intellectual  | 6.7              | 32.1        | <b>33.6</b> | 27.6  | 100.0 | 134   |                |
|                | total         | 21.2             | <b>32.6</b> | 28.1        | 18.1  |       | N 420 |                |

WT types are connected to ST performance processes ( $\chi^2=26.0$ ,  $p=.002$ , Table 10). The trends are not as clear as among information needs. Single processes are the largest class in all WT types, while their share is the smallest, 36%, in intellectual tasks. Interestingly, in support tasks, developing processes are almost as common (22%) as in intellectual tasks (25%). List processes form a fourth of STs in intellectual tasks.

TABLE 10. Search process types (%) in WT types.

|                |               | search task performance process |      |        |            |       | total | n search tasks |
|----------------|---------------|---------------------------------|------|--------|------------|-------|-------|----------------|
|                |               | single                          | list | stable | developing |       |       |                |
| work task type | communication | <b>60.6</b>                     | 12.1 | 15.2   | 12.1       | 100.0 | 132   |                |
|                | support       | <b>47.3</b>                     | 16.4 | 14.5   | 21.8       | 100.0 | 55    |                |
|                | editing       | <b>60.6</b>                     | 17.2 | 10.1   | 12.1       | 100.0 | 99    |                |
|                | intellectual  | <b>35.8</b>                     | 24.6 | 14.9   | 24.6       | 99.9  | 134   |                |
|                | total         | <b>51.0</b>                     | 17.9 | 13.8   | 17.4       |       | N 420 |                |

### *The combined effect of task type and complexity*

In this section, we analyze the combined effects of WT types and WT complexity on STs. Here, we classify WT complexity in two classes, simple and complex, by combining the complexity classes I-II, and III-IV (Table 3). Correlation coefficients are calculated between (continuous) task complexity and suitable ST variables within each WT type. Table 11 presents collectively the figures for all WT types, and the results are discussed by WT type.

TABLE 11. WT complexity's effects on STs within WT types.

|                       |                      | WT type       |         |         |         |         |         |              |         | total |
|-----------------------|----------------------|---------------|---------|---------|---------|---------|---------|--------------|---------|-------|
| ST feature            | category             | Communication |         | Support |         | Editing |         | Intellectual |         |       |
|                       |                      | simple        | complex | simple  | complex | simple  | complex | simple       | complex |       |
| main search medium(%) | <i>pc</i>            | 2.2           | 0.0     | 6.7     | 2.5     | 0.0     | 4.7     | 3.3          | 5.8     | 2.9   |
|                       | <i>org_database</i>  | 22.2          | 41.4    | 20.0    | 30.0    | 41.1    | 23.3    | 30.0         | 34.6    | 33.1  |
|                       | <i>Web</i>           | 31.1          | 14.9    | 53.3    | 40.0    | 41.1    | 27.9    | 30.0         | 45.2    | 33.8  |
|                       | <i>communication</i> | 44.4          | 43.7    | 20.0    | 27.5    | 17.9    | 44.2    | 36.7         | 14.4    | 30.2  |
|                       | <i>total</i>         | 99.9          | 100.0   | 100.0   | 100.0   | 100.1   | 100.1   | 100.0        | 100.0   | N 420 |

|   |                          |       |       |       |       |       |       |       |       |       |
|---|--------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| <b>no. of unique search media information need(%)</b> | <i>average</i>           | 1.07  | 1.05  | 1.27  | 1.18  | 1.16  | 1.09  | 1.20  | 1.38  | 1.18  |
|   | <i>instrumental fact</i> | 26.7  | 34.5  | 33.3  | 27.5  | 28.6  | 14.0  | 20.0  | 2.9   | 21.2  |
|   | <i>known item</i>        | 31.1  | 25.3  | 26.7  | 45.0  | 39.3  | 32.6  | 30.0  | 32.7  | 32.6  |
|   | <i>topic</i>             | 22.2  | 27.6  | 33.3  | 17.5  | 19.6  | 37.2  | 33.3  | 33.7  | 28.1  |
|   | <b>total</b>             | 20.0  | 12.6  | 6.7   | 10.0  | 12.5  | 16.3  | 16.7  | 30.8  | 18.1  |
| <b>performance process(%)</b>                         | <b>total</b>             | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.1 | 100.0 | 100.1 | N 420 |
|   | <i>single</i>            | 64.4  | 58.6  | 53.3  | 45.0  | 64.3  | 55.8  | 26.7  | 38.5  | 51.0  |
|   | <i>list</i>              | 8.9   | 13.8  | 13.3  | 17.5  | 14.3  | 20.9  | 30.0  | 23.1  | 17.9  |
|   | <i>stable</i>            | 13.3  | 16.1  | 20.0  | 12.5  | 12.5  | 7.0   | 23.3  | 12.5  | 13.8  |
|   | <i>developing</i>        | 13.3  | 11.5  | 13.3  | 25.0  | 8.9   | 16.3  | 20.0  | 26.0  | 17.4  |
|   | <b>total</b>             | 99.9  | 100.0 | 99.9  | 100.0 | 100.0 | 100.0 | 100.0 | 100.1 | N 420 |
|   | <i>n of STs</i>          | 45    | 87    | 15    | 40    | 56    | 43    | 30    | 104   | N 420 |

### Communication tasks

In complex communication tasks, organizational databases are much more common than in simple tasks. In the case of Web use the trend is clearly the opposite. Communication media are used equally often in simple and complex tasks. The number of unique search media, the types of information needs and ST processes respond weakly to WT complexity.

### Support tasks

The use of organizational databases and the Web clearly changes with WT complexity in support tasks: The Web is the most common main search medium in simple and in complex tasks, but its share decreases whereas the share of organizational databases increases with growing task complexity. Interestingly, fewer search media are used in complex than in simple WTs. A similar trend can be seen in editing tasks.

In support tasks, WT complexity is negatively correlated with information need complexity ( $q=-.31, p=.021$ ). Factual needs grow remarkably when WTs become complex: their share increases from 27 % to 45%. The need for known items decreases from 33% to 18%. Editing tasks have the opposite trend with known items.

In complex support tasks, the share of developing searches is as large as in complex intellectual tasks (a quarter), which is remarkably larger than in other task types. The share of developing searches is *clearly* larger in complex than in simple tasks; this is unique to support tasks.

### Editing tasks

The shares of main search media clearly change with WT complexity in editing tasks. The use of organizational databases and the Web together in simple editing tasks decreases from over 80% to about 50% in complex tasks. Communication media become prominent in complex tasks: their share increases from 18% to 44%.

In editing tasks the need for known items differs between simple and complex tasks - but in the opposite direction compared to support tasks: in simple tasks, known items form 20% of information needs but 37% in complex tasks. Instrumental information needs decrease from 29% to 14%. WT complexity is positively correlated with information need complexity ( $q=.24, p=.015$ ).

### Intellectual tasks

In intellectual tasks, the use of the Web increases clearly with growing task complexity, whereas the use of communication media decreases. Unlike other task types, complex intellectual tasks include more diverse search media use than simple ones. The share of STs with more than one medium increases from 13% to 28%.

Intellectual tasks show statistically significant connections ( $\chi^2=11.9, p=.008$ ) between WT complexity and the distribution of information needs. The shares of facts and known items remain the same, the share of topical needs almost doubles, and the share of instrumental needs drops from 20% to 3% with growing WT complexity. This shows also in the correlation between information need complexity and WT complexity ( $q=.23, p=.007$ ): the more complex the

task, the more complex information needed.

Complex intellectual tasks have an exceptionally large share (26%) of developing search processes. However, the share of single searches is actually bigger in complex than in simple tasks. The difference is exceptionally clear compared to other task types. The share of stable search processes decreases from simple to complex tasks.

## Discussion

### *WT complexity affects searching*

The main search medium and the number of unique search media were quite independent of WT complexity. The most complex WTs, however, had more search media than other tasks. Earlier studies of information seeking (e.g. Byström & Järvelin, 1995; Saastamoinen et al., 2012) show that task complexity affects information source use, but WT complexity seems not to affect the search media used.

A weak correlation suggested that the more complex the WT, the less flexible was the search medium used for searching. This finding suggests that each search medium has varying roles in WTs and thus the flexibility of the information offered is not a good defining feature in this data. Again, search media differ from information sources in this respect, because previous studies (e.g. Byström & Järvelin, 1995) rather show that the more complex the task, the more flexible information sources used.

The share of topical needs peaked in complex tasks. This finding seems straightforward and supports the finding by Saastamoinen, Kumpulainen, Vakkari, and Järvelin (2013). However, the share of topical needs in the most complex tasks is only a quarter; simpler information needs are still important.

In our data, single search processes were most common overall. Exploratory, that is developing processes were most common in the most complex tasks. Our findings partly support Kumpulainen's (2014) conclusion that certain types of processes (trails) are connected to certain task complexity categories.

### *WT types affect searching*

WT type was surprisingly strongly connected to ST features. As the clearest example, all information needs have linear trends along *sup-edit-int* continuum. The shares of topical needs and known items grow, and the shares of factual and instrumental needs decrease following the trends that Saastamoinen and colleagues (2013) found.

Participants across varying work roles and information professions seem to identify *communication* as a separable task type that has to be taken care of on a daily or continuous basis. Communication is normally supporting other tasks (e.g. no emails written *per se*) but considered more than a mere support task. Table 3 suggests that communication tasks are between intellectual and editing tasks regarding their complexity.

### *Task types versus complexity*

We suggest that task type and complexity augment each other and should be analyzed together. In the present study, task complexity was calculated as an average of numerical estimates given by the participants, whereas task types were classified by the researcher based on the written, open task descriptions. Therefore WT complexity was rather subjective, whereas WT type was somewhat more objective. At least task types were selected using the same criteria for all tasks but each participant may have had their own understanding of complexity criteria.

Growing task complexity seems to increase the need for new information, whereas task type seems to be connected to the type of the information needed: in support tasks, factual needs grow rapidly from simple tasks to complex tasks, whereas the same happens to topical needs in intellectual tasks. However, the apparent complexity of the STs does not necessarily correspond to the complexity or type of the underlying WTs.

Without a doubt, in communication tasks, communication media are as important regardless of task complexity. In support and editing tasks, communication media are used more in complex than in simple tasks, which may reflect the need for flexible information. Surprisingly this does not hold true for complex intellectual tasks where communication tools are rarely the main search media. It may be that complex intellectual tasks must be performed individually and consulting other people would not help – especially in case of complex tasks – as intellectual tasks are at the core of one's own expertise. Intellectual tasks in our data were mainly writing tasks where the participant was the expert herself; and not extensive decision tasks that have been of much interest in earlier studies (see for example the review

by Li and Belkin (2008)). These findings may depend on the work roles of the participants. On the other hand, we did not analyze non-computer mediated information interaction which may have taken place in all task types; perhaps performing complex intellectual tasks demanded face-to-face support from colleagues.

### *About real-life searching*

Our point of departure was that since WTs are building blocks of information-intensive work, they must also affect STs, an integral part of these WTs. Because real-life studies on this scale are rare, there was no basis for hypothesis formulation and testing. Both methodologically and research question-wise the present study belongs to the same category as the studies by Hansen (2011), Saastamoinen et al. (2012) and Kumpulainen and Järvelin (2010).

STs were mainly simple: the majority consisted of a single query in a single search medium while the media differed between tasks. Ingwersen and Järvelin (2005) postulated that simple, organized practices supply most of the information needed in work. It seems important to develop systems to serve also simple STs effectively and efficiently, but the tasks in IIR experiments typically represent complex, search-intensive tasks. Often they are not even WTs but related to leisure time or other non-work issues. To start with, further studies are needed to find out how more typical WT scenarios could be exploited in IIR research.

## **Limitations**

The present study was an explorative real-life study, and its data set was rich but initially rather unstructured. It had to be processed manually in many phases, and categorizations and decisions were made intellectually from the start. The decisions were based on all available data, including the transaction and video logs.

Since we wanted to maximize the validity of the study, we used as much of the data as possible; that is, chunks of data were discarded only for strong reasons. This approach differs from laboratory studies and other studies conducted in controlled environments where the internal integrity of data is highly valued and only perfect logs, for example, are accepted for further analysis. Some data was missed if the participants occasionally forgot to start the recording or to save it; or wanted videos to be blurred or cut for confidentiality reasons.

However, it was possible to reliably reconstruct the missing parts of the searches. This was possible because there was plenty of data regarding similar search situations. We could also triangulate the missing data using the available data subsets. We analyzed the distribution of the augmented data and how it affected the results, and nothing questioned our approach of augmenting the data.

Our study was rather ethnographic, which challenges its repeatability. This is the price one must pay when studying humans in their natural environments. The risk is worth taking if one wants to understand these phenomena better in order to design new kinds of more structured studies and to innovate new kinds of information systems.

It is possible that there are other factors that affect information searching indirectly besides WT complexity and WT type. Our data suggests that task complexity distributions may vary by organization type. Further analysis of the effects of organization or other contextual factors was beyond our research questions and is left to future research.

## **Conclusions**

In this paper, authentic search tasks (STs) were empirically studied in the context of authentic work tasks (WTs). This type of research is rare in the field of information seeking and retrieval. ST - WT interactions were analyzed statistically using information needs, ST processes and search media as dependent variables, and WT type and complexity as independent variables.

It was found that WT type may have clearer connections to STs than WT complexity, and that most STs are simple, i.e., including only one query in one search medium despite WT complexity or type. The studied authentic WTs do not much resemble the so called simulated WTs used in past studies. Similarly, neither do the authentic STs much resemble test tasks applied in user studies in interactive information retrieval because the information retrieval part of authentic work is inseparable from its other subtasks.

The present study may help assess the applicability of experimental STs used in future research. The main question for the future is whether information systems and information retrieval features are to be developed to support all kinds of STs/WTs or only some task types. This includes the questions of whether to concentrate on complex/intellectual WTs where information needs are more vague, or to include other kinds of tasks, as well.

This paper contributes to the knowledge of authentic information searching. The findings can be used to design more realistic experiments in future, to better understand what information searching is about in real-life contexts, and therefore guarantee validity of experimental information search studies.

## Acknowledgment

The authors wish to thank Ms. Heljä Franssila (University of Tampere) for her cooperation.

## References

- Algon, J. (1997). Classifications of tasks, steps, and information-related behaviors of individuals on project teams. In P. Vakkari, R. Savolainen, & B. Dervin (Eds.), *Information seeking in context. Proceedings of an international conference on research in information needs, seeking and use in different contexts* (pp. 205-221). London, UK: Taylor Graham Publishing.
- Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review*, 13, 407-424.
- Borlund, P. (2003). The IIR evaluation model: A framework for evaluation of interactive information retrieval systems. *Information Research*, 8, paper 152.
- Broder, A. (2002). A taxonomy of web search. *ACM Sigir Forum*, 36, 3-10.
- Byström, K., & Hansen, P. (2005). Conceptual framework for tasks in information studies. *Journal of the American Society for Information Science and Technology*, 56, 1050-1061.
- Byström, K., & Järvelin, K. (1995). Task complexity affects information seeking and use. *Information Processing & Management*, 31, 191-213.
- Du, J. T. (2014). The information journey of marketing professionals: Incorporating WT-driven information seeking, information judgments, information use, and information sharing. *Journal of the Association for Information Science and Technology*, 65, 1850-1859.
- Foster, J. (2006). Collaborative information seeking and retrieval. *Annual Review of Information Science and Technology*, 40, 329-356.
- Gwizdka, J., & Spence, I. (2006). What can searching behavior tell us about the difficulty of information tasks? A study of web navigation. *Proceedings of the American Society for Information Science and Technology*, 43, 1-22.
- Hackman, J. R. (1968). Effects of task characteristics on group products. *Journal of Experimental Social Psychology*, 4, 162-187.
- Hansen, P. (2011). Task-based information seeking and retrieval in the patent domain. Processes and relationships. Academic Dissertation, University of Tampere.
- Ingwersen, P. (1986). Cognitive analysis and the role of the intermediary in information retrieval. In R. Davies, (Ed.) *Intelligent information systems* (pp. 206-237). Chichester, UK: Horwood.
- Ingwersen, P., & Järvelin, K. (2005). *The turn: Integration of information seeking and retrieval in context (the information retrieval series)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc.
- Järvelin, K., Vakkari, P., Arvola, P., Baskaya, F., Järvelin, A., Kekäläinen, J., . . . Sormunen, E. (2015). Task-based information interaction evaluation: The viewpoint of program theory. *ACM Transactions on Information Systems*, 33(1), 3:1-3:30.
- Kellar, M., Watters, C., & Shepherd, M. (2006). The impact of task on the usage of web browser navigation mechanisms. *Proceedings of Graphics Interface 2006* (pp. 235-242). Toronto, Ont., Canada: Canadian Information Processing Society.
- Kelly, D. (2006). Measuring online information seeking context, part 2: Findings and discussion. *Journal of the American Society for Information Science and Technology*, 57, 1862-1874.
- Kelly, D., Arguello, J., & Capra, R. (2013). NSF workshop on task-based information search systems. *SIGIR Forum*,



47, 116-127.

Kotov, A., Bennett, P. N., White, R. W., Dumais, S. T., & Teevan, J. (2011). Modeling and analysis of cross-session search tasks. *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval Beijing, China*, (pp. 5-14). New York, NY, USA: ACM.

Kumpulainen, S. (2014). Trails across the heterogeneous information environment. *Journal of Documentation*, 70, 856-877.

Kumpulainen, S., & Järvelin, K. (2010). Information interaction in molecular medicine: Integrated use of multiple channels. *Proceedings of the Third Symposium on Information Interaction in Context, New Brunswick, New Jersey, USA* (pp. 95-104). New York, NY, USA: ACM.

Li, Y. (2009). Exploring the relationships between work task and search task in information search. *Journal of the American Society for Information Science and Technology*, 60, 275-291.

Li, Y., & Belkin, N. J. (2008). A faceted approach to conceptualizing tasks in information seeking. *Information Processing & Management*, 44, 1822-1837.

Liu, J., Kim, C. S., & Creel, C. (2015). Exploring search task difficulty reasons in different task types and user knowledge groups. *Information Processing & Management*, 51, 273-285.

Liu, P., & Li, Z. (2012). Task complexity: a review and conceptualization framework. *International Journal of Industrial Ergonomics*, 42, 553-568.

Maynard, D. C., & Hake, M. D. (1997). Effects of objective and subjective task complexity on performance. *Human Performance*, 10, 303-330.

Nicholas, D., Williams, P., Rowlands, I. & Jamali, H.R. (2010). Researchers' e-journal use and information seeking behaviour. *Journal of Information Science*, 36, 494-516.

Pharo, N. (2004). A new model of information behaviour based on the search situation transition schema. *Information Research*, 10, paper 203.

Pharo, N., & Järvelin, K. (2004). The SST method: A tool for analysing web information search processes. *Information Processing & Management*, 40, 633-654.

Saastamoinen, M., & Järvelin, K. (2016). Queries in authentic work tasks: The effects of task type and complexity, 72, 1114-1133.

Saastamoinen, M., Kumpulainen, S., & Järvelin, K. (2012). Task complexity and information searching in administrative tasks revisited. *Proceedings of the 4th Information Interaction in Context Symposium, Nijmegen, The Netherlands* (pp. 204-213). New York, NY, USA: ACM.

Saastamoinen, M., Kumpulainen, S., Vakkari, P., & Järvelin, K. (2013). Task complexity affects information use: A questionnaire study in city administration. *Information Research*, 18(4), paper 592.

Toms, E.G. (2011). Task-based information searching and retrieval. In I. Ruthven and D. Kelly. (Eds.) *Interactive information seeking, behaviour and retrieval* (pp. 43-59). London: Facet Publishing.

Vakkari, P. (2003). Task-based information searching. *Annual Review of Information Science and Technology*, 37, 413-464.

Vakkari, P., Pennanen, M., & Serola, S. (2003). Changes of search terms and tactics while writing a research proposal A longitudinal case study. *Information Processing & Management*, 39, 445-463.

Wildemuth, B., Freund, L., & Toms, E. G. (2014). Untangling ST complexity and difficulty in the context of interactive information retrieval studies. *Journal of Documentation*, 70.

## Appendix 1: Morning questionnaire

*What WTs do you have to perform today? Base your replies on your present knowledge.*

### Contact information

Name:

Organization:

### Tasks

[The questionnaire includes place holders for 5 tasks. The participants fill in as many tasks as they need to. For each:]

*The WTs of today. Task complexity should be given as a number 0 (really simple)-100 (really complex).*

Task description:

Task complexity 0%-100%:

*Task performance. You should answer with a number 0 (not at all)-100 (perfectly) to the question of how well you know the task performance process.*

Programs/information systems needed:

Other information/sources of information needed:

How well do you know the task performance process?:

Other remarks (e.g., about the tasks or the course of the day):

## Appendix 2: Afternoon questionnaire

*What WTs did you perform today? Think about the whole working day from the morning to the afternoon as you reply.*

### Contact information

Name:

Organization:

### Tasks

*The WTs of today. Task complexity should be given as a number 0 (really simple)-100 (really complex).*

[The questionnaire includes place holders for 5 tasks. The participants fill in as many tasks as they need to. For each:]

Task description:

Task complexity 0%-100%:

At what time (approximately) did you perform the task? (E.g., 12-13)

*Task performance*

Programs/information systems you used:

Other information/sources of information you used:

Did any problems turn up in the task? Please describe them.

Other remarks (e.g. about the tasks or the course of the day):