Article

JIS

Journal of Information Science
2018, Vol. 44(2) 265–284

DOI: 10.1177/0165551516687726
journals.sagepub.com/home/jis
SAGE

# Relationships between work task types, complexity and dwell time of information resources

**Miamaria Saastamoinen**
Faculty of Communication Sciences, University of Tampere, Finland

**Kalervo Järvelin**
Faculty of Communication Sciences, University of Tampere, Finland

**Abstract**
Information seeking research often reports about types of information resources, ways of acquiring them and opinions on their importance in various professions. Based on self-reporting, these findings are affected by human memory and rationalisation. This article proposes a new way of studying information resource use – based on dwell time in the context provided by concrete work tasks. We use log data of 21 information workers from six organisations to analyse how work task complexity is connected to the time used in various information resources; how task complexity is connected to information resource use in different task types. Unlike traditionally, our findings consist of objective data on which resource types are used, and for how long, in work tasks of varying complexity and type. For example, the findings suggest that growing work task complexity increases the dwell time in local personal computer (PC) resources; these resources are especially popular in intellectual tasks. Such findings help understand factors affecting information resource use. Likewise, they help focus attention on most time-consuming aspects of task-based information interaction when developing support for work.

**Keywords**
Field studies; information resources; information seeking knowledge work; task complexity; transaction logs; work tasks

## 1. Introduction

Information seeking is a phenomenon that can be defined as the process where people are trying to fulfil their information needs using information resources. It has inspired several models [1] that describe the factors that can affect the seeking process. One of the most crucial factors is the information seeker's task [2]. Especially, task complexity has had a growing interest among researchers [3–5]. People perform tasks in their leisure time as well, but a task is perhaps a more easily separable part of one's work which is further affected by the context of the task, for example, the organisation in question. In this article, we analyse information resource use in work tasks observed in real time. The analysis goes beyond mere seeking, including all kinds of information interaction [6]. We have earlier proposed a task type categorisation [7,8]. This categorisation is abstract and data-driven and can be used to elaborate the effects of work task complexity on information interaction.

Our mission is to describe real-life information interaction in the context of work tasks. The connections between information *seeking* and tasks are often suggested in the literature but seldom analysed in real-life settings using real-time methods. The research questions are as follows:

- What are the overall differences in computer-based information resource use between various task complexity and task type categories?

**Corresponding author:**
Kalervo Järvelin, Faculty of Communication Sciences, University of Tampere, Tampere 33014, Finland.
Email: kalervo.jarvelin@staff.uta.fi

- How does task complexity affect computer-based information resource use within task types?

This study is a unique combination of research methods, objectives and findings. The article contributes to the field of information interaction by

- Knocking down barriers between different schools of thought;
- Proposing new combinations of data collection and analysis methods;
- Presenting new empirical findings to be used by both practitioners and researchers;
- Suggesting new questions for future research.

Next, we will present some prevailing confrontations and gaps within information behaviour research and how they can be overcome.

*Information seeking and information retrieval* stem from two different schools of thought. Nowadays, electronic information (retrieval) systems clearly play an important role in both leisure time and organisational life, but information seeking and retrieval are still often considered as two different objects of research. Lowering the boundaries between these two was already called for by Kuhlthau [9], Järvelin and Ingwersen [10] and Vakkari [11]. Our study contributes to their unification by considering information retrieval systems, that is, search engines, as an equal computer-based information resource among others, such as locally saved files and organisational databases. The use of all computer-based resources is analysed in the same way, that is, by calculating the time spent using each resource type (dwell time).

*Qualitative and quantitative methods* are often seen as competing or opposite to each other though they actually belong to the same continuum [12]. Our study starts with a qualitative, real-life data set that is analysed in detail to provide quantitative data about connections between work tasks and information interaction.

Analysing Web searching behaviour using server-side logs is common [13]. The data are readily quantitative and typically large, enabling powerful statistical analyses. These logs, however, lack explicit information about the situation the searcher was in when searching. A way to gain insight into the context is to use self-report methods [3,14–18] which can however be unreliable. To overcome these problems, we use several data collection methods, including direct observation, interviews, questionnaires, client-side logs and screen video capture.

In practice, *working day is the time* that is allocated to various tasks and resources. Dwell time is a popular variable in interactive information retrieval (IIR) research since it has been widely found to indicate interest and document relevance [19–21]. In these studies, dwell time is understood in a narrow way, as the time spent *reading* USENET news [21], articles related to studying [20] or web pages [22]. Some researchers have suggested that dwell time does not self-evidently reflect interest but depends also on task and prior knowledge [22–24].

We broaden the earlier dwell time approach by analysing dwell time in all resources the participants use, such as email, web pages and local files. We study *all kinds of resources* for information interaction instead of only sources used to find new information; information interaction in work tasks is a complete process where searching for and selecting, using and synthesising information intertwine [6]. First, we analyse what these resources are (qualitative analysis) and then their use quantitatively as dwell time, that is, time spent using each resource in a task. However, we only analyse information resource use taking place through the study participants' computers, not any manual actions. While this is admittedly a limitation, given the growing digitalisation of work, our study sheds new light on this growing share of information interaction.

Our participants include various professional knowledge workers; a typical approach in studying information interaction is to focus solely on university students performing pre-designed search or work tasks [25,26]. All work tasks and all information interactions in our data set are authentic field data with no control by the researchers.

In order to advance our knowledge on information interaction and to design better information systems and information environments, we should aim at understanding information interaction as a part of authentic work tasks. This is the mission of this article.

This article belongs to a series of papers written by the authors of this article and analysing the same large data set with different research questions. Our earlier article [7] includes the analysis of search tasks (i.e. subtasks of work tasks conducted by querying in information systems) and article [8] includes detailed analysis of queries in real-life data.

This article is structured as follows. In section 'Literature', we first discuss tasks as context for information seeking, especially task complexity and task type. This is followed by a discussion about information resources in earlier research. Section 'Methods' describes study methods in detail: the participants, data collection and analysis methods, and work task and information resource categories used. Section 'Results' includes detailed findings, first by work task complexity, second by work task types and third by work task complexity within each work task type. The findings are discussed on a more general level in section 'Discussion'. This is followed by sections 'Limitations and implications' and 'Conclusion'.

## 2. Literature

In this section, we discuss earlier research on tasks and information seeking. First, the theoretical ideas of work tasks and task complexity are presented, and second, some important empirical studies that use work tasks and especially task complexity as independent variables are presented. The section ends with discussion about information resources. Please note that our empirical findings concern only *work* tasks as independent variables, not, for example, leisure-time tasks or search tasks that are subtasks of work tasks. However, the concept of task is discussed below on a more general level, following the literature; differences between task conceptions are made explicit when necessary.

### *2.1. Tasks as context*

Courtright [27] argues that the research into information behaviour does not sufficiently focus on the theoretical concept of context. Context is often mentioned but left as a rather stable background of whatever is the object of interest. Courtright [27] lists tasks, problems and situations as one commonly used understanding of context. In our study, tasks are a context in the sense that they include and affect information interaction. They form a clearly defined empirical variable thus actually being in focus, not outside of it. The organisation acts as a backdrop that is not used for analysis. This type of understanding of tasks works well as conceptualisation of context: Courtright [27] calls for context that is seen as flexible and more of a doing than being. Subjective work tasks are the participants', the actors', constructed situations that they must react to. The actor herself can define how much the task is shaped by the social situation or organisation and how much of her own personal mindset. Context is continuous, pervasive interplay between everything [27].

Many studies have found tasks to be an important context influencing information-related behaviour [28]. It is natural since information seeking, for instance, depends on the larger task it contributes to Järvelin and Ingwersen [10]. For example, Li's [14] study shows several connections between work tasks and search tasks. Thus, it is crucial to include the work task aspect in studies; analysing authentic work tasks is especially beneficial since the task performance is probably to be authentic if the task has real consequences for its performer [29]. Following these basic ideas, our point of departure is the participants' work tasks that trigger information needs and use of information resources. As Kim and Soergel [30] suggest, the best way to understand task-based information behaviour is to study authentic tasks by authentic performers in authentic situations.

Tasks can be analysed on several levels. Byström and Hansen [31] differentiate between work tasks, information seeking tasks and information search tasks (information retrieval tasks). They also differentiate between abstract task descriptions and tasks as processes. In this study, the unit of analysis is work task, and we do not analytically identify their components such as seeking tasks. This would have been possible but is out of our scope since we aim to study the larger context's (i.e. the work tasks') effect on information resource use. Furthermore, we discuss only work tasks that manifest in the logs, that is, 'tasks as processes' as opposed to abstract tasks. However, these concrete work tasks have features (task complexity and task type) that enable comparisons between tasks.

Several studies have operationalised larger tasks as tasks related to leisure time or studying, and participants are typically students performing authentic or simulated tasks. These include Li and Belkin's [32] experiment on students' information searching in several systems, Pharo's [33] realistic study of students' course work including searching, Taylor's [34] similar analysis of information retrieval behaviour of students making a slideshow and Toms et al.'s [35] study of mainly students performing leisure-time-related simulated tasks. Although these studies definitely provide us with information on how larger tasks affect information searching, their focus on students and tasks related to studying or simulated leisure-time interests may be limited in applicability. Also, the available information systems and resources may differ between work and leisure contexts, the former providing several task-specific systems/resources. Our approach to work tasks aims at expanding the knowledge of work task performance to cover actual, authentic work tasks that are observed using real-time methods in authentic environments. At least we should have a clear understanding of authentic work tasks in order to reliably simulate their performance in controlled settings. Consequently, interpretation of the findings at least partly depends on how the data are collected: whether using simulated or authentic tasks, and to what extent the types of tasks are applicable outside the studied situation.

*Task complexity* is widely recognised as a key factor explaining various information activities, but the field lacks a coherent understanding of 'complexity'. Liu and Li [36] make a significant effort to map the concept of task complexity. According to Liu and Li [36], the concept of task complexity is perplexed especially by the unclear differences between, first, objective and subjective complexity and, second, between task complexity and task difficulty. Liu and Li [36] try to connect various studies: subjective task complexity and task difficulty are similar in the sense that they are often considered as dependent on task performer and her experience. Task complexity per se, as opposed to task difficulty, then refers to the objective features of the task itself as an abstraction, rather than the features of the performer [36]. Maynard and Hakel [26] discuss, based on an empirical test setting, findings on the interrelationships between objective and subjective task complexity. Wildemuth et al. [5] discuss search task difficulty versus complexity in their review: these are affected by, for example, number of facets and subtasks. Their notions are applicable to the information seeking as part of work task performance, as well. Each study must operationalise and define task concepts suitably for its research questions; however, for example, the differences between task complexity and difficulty may be hard to separate and measure accurately in empirical studies even though separable in theory.

Earlier studies do discuss and analyse subjective task complexity as presented above, but it is sometimes considered inferior to objective task complexity that may seem easier to measure and understand. However, we argue that subjective work task complexity is a beneficial variable when analysing authentic work tasks. Without a doubt, tasks have features that relate to the objective complexity, but these are hard to measure in real-life settings if the researcher is not a specialist on each substance field studied. The task performer herself is nearest to the task and it is hard to imagine that her task performance would not be affected by her situational task complexity expectation.

Interestingly, if we assume that search (or seeking) tasks are part of work tasks and these subtasks vary in their complexity, we come to a situation where the effects of complexity work in both ways. The complexity of work task affects information seeking, and the complexity of seeking affects the complexity of the work task. In this article, we discuss the phenomenon only one-way, that is, how work task complexity affects information resource use. However, the interaction of these phenomena should be kept in mind. Logic requires that if any subtasks are complex, they increase the complexity of the work task. Thus, if information seeking is considered complex, the task performer may feel the whole work task complex (or difficult) even before starting it, and if information seeking proves complex, it certainly affects her estimate of complexity afterwards.

Next, we discuss previous studies analysing the effects of work task complexity and type. We focus on studies that include task categories similar to the ones analysed in this article.

### 2.2. Work task complexity and type as empirical variables

In empirical studies, task complexity has been found to affect task performance and information seeking. Byström [3] and Kumpulainen and Järvelin [37] operationalised task complexity as the amount of prior knowledge concerning each task. Byström [3] measured it using daily responses of participants, Kumpulainen and Järvelin [37] estimated the complexity based on how well the task performer was able to describe the task beforehand. Byström [3] found that the more complex the task, the more human resources and the less documentary resources are used. Kumpulainen and Järvelin [37] found that if task performers have good prior knowledge about their work tasks, the performance process is more straightforward than in tasks with less prior knowledge.

Our study closely follows the understanding of task complexity and the overall approach to authentic work tasks in the studies by Byström [3] and Kumpulainen and Järvelin [37]. Byström's [3] pioneering work provided us suitable questions to pose to the participants concerning work task complexity, that is, the knowledge of task process and direct question about work task complexity. The participants in the study by Byström [3] were asked to draw their estimates on pictures that provided the scale from 0 to 100. We used electronic questionnaires where the participants gave their numerical estimates that are easier to analyse further since they are exact. Byström [3] had also other measures of complexity; one major difference is that she did not use the participants' complexity estimates directly as we did, but formed an intellectual aggregate of them and used case-specific consideration when assigning work tasks to complexity categories. Kumpulainen and Järvelin [37] based their complexity estimates on face-to-face questions to participants but as researchers decided whether the participant knew the task process, the information resources needed and the task outcome. We had data collection days also outside direct observation, and thus, this procedure was not feasible. Our earlier experience with collecting data about work task complexity [38] helped us select the three complexity measures used in this article. Therefore, theoretically our concept of task complexity is similar to Byström's [3] and Kumpulainen and Järvelin's [37] while its operationalisation is different.

Tasks have also been categorised by other task features than task complexity in the literature. Li and Belkin [2] made a thorough review of task types. Du et al. [39] studied marketing professionals and formed a categorisation of their tasks including administrative tasks and writing reports. The categorisation was tailor-made though easily applicable [39]. Hansen [40] studied the information seeking processes of patent engineers and used different types of patent applications and different phases of patent handling as task types. Highly specialised, data-driven task types are unfortunately difficult to compare with other studies that analyse tasks of other fields. Thus, we have followed a more abstract way of categorising work tasks, compared with, for example, Hansen [40]. Our task types are presented in detail in section 'Defining task types and complexity'.

### 2.3. Information resources

We decided to use the term information *resource* instead of information *source*, because in realistic task performance, different aspects of information interaction such as creating new information, searching for information and using information intertwine. These are actually difficult to distinguish in log data. Furthermore, the separation seems purposeless considering the aims of the present research. Thus, information resources cover information sources. Some categories are more probably to act as pure sources (e.g. search engines), while some are more potential resources for creating information (e.g. word processing software). Saastamoinen and Kumpulainen [41] mapped the features of information resources and resource categorisations applied in information seeking studies.

Our study discusses only the use of computer-based information resources. In the organisations studied here, and perhaps in knowledge work more generally as well, using computer-based information resources takes a large part of the working day in terms of time, and paper-based repositories of information are not used much since information is managed in digital form within organisations. Furthermore, studying the dwell times of non-computer-based information resources would require someone to observing the participants' work continuously, ready for taking time with a stopwatch. That would be extremely distracting for the task performer. We discuss resource categories that are similar to ours and used in earlier studies below.

*Communication* is an information resource typically included in information seeking studies. Savolainen [42] names human resources such as colleagues and experts but still considers email as a networked, not human, resource. Byström [3] found that people are an important resource for municipal officers since they use human resources to find all types of information. She also found that the use of human resources clearly increases with growing task complexity [3]. Hertzum's [43] review deals extensively with human resources, and he shows that human resources often belong to the most used resources. Several studies also show that task complexity is connected to the increased use of human resources [43]. Having a communication resource category seemed thus necessary in terms of earlier research, and also our empirical data show that communication resources are widely used in (computer-mediated) knowledge work.

The use of *information retrieval systems* is of course in focus in information retrieval studies, but it is also taken into account as an information resource among others in some studies. The electronic information environments nowadays support the idea of including search engines as a major category of information resources. Kumpulainen and Järvelin [37] studied the information resources used by researchers using workflow charts that represented typical processes of work tasks. Their resource classification included Web search engines, but they also took notice of queries in other systems [37]. Nicholas et al. [44] studied e-journal use and they found that academics often use various search engines to start searching for relevant papers. Similarly, Du [16] found that search engines are widely used in the work of marketing professionals. Depending on the study, especially general-purpose search engines may be considered a category of its own or part of the *World Wide Web*. Hansen [40] mentions web pages as one information resource type of patent engineers, and Savolainen [42] identifies 'networked resources', including web pages, as a resource for problem-specific information for non-work purposes. Our data supports the view of the Web as an information resource that is used by several professionals across different fields, and thus, it is a category of its own. As some earlier researchers, we also separated search engine use from other Web use because they are used in different ways and for different purposes. Furthermore, the separation makes the findings more comparable to studies of IIR.

When work-related information seeking is studied, it is quite convenient to include *organisation-specific resources* into the study. Byström's [3] study does not have this as an explicit category, but her categories of official documents and registers can be considered as such. Her findings show that the use of official documents decreases with growing task complexity [3]. Du [16] found that internal databases are a frequently used resource of marketing professionals. In addition to special software and internal databases, especially large organisations often have an intranet in use. It is a heterogeneous collection of official documents, instructions, useful external links and announcements that need to be available but do not fit elsewhere. Today's organisational information environment in knowledge work clearly takes advantage of internal information systems where all important information can be gathered in a rather efficient and accessible way. Thus, this category is essential when studying authentic knowledge work.

The information resources on *personal computer* (PC) or other personal collections are seldom discussed in earlier literature. However, 'personal files' were discussed already in the model by Taylor [45] where they were seen as the alternative to using a library. PC as a resource category is also taken into account in an empirical study by Kumpulainen and Järvelin [37] as an important resource alongside field-specific (molecular medicine) databases. Kumpulainen and Järvelin [37] show that, for example, reformulating queries using software on a PC may be an important part of the information seeking process. Following the findings in Kumpulainen and Järvelin [37], our earlier work [38] and the trends in the present data, PC is an important resource to be included in the analysis.

## 3. Methods

### 3.1. Participants

The participants are knowledge workers who were contacted directly or via their organisation using the authors' existing contacts to ease finding enough participants. We selected participants who were willing and able to participate and had enough tasks that were performed using a computer. The first author performed the data collection. The logs were collected during the period from August 2013 to October 2014. Table 1 includes participant details.

The organisations and participants form a convenience sample. Finding participants was not a question of selecting them from a large pool of ready volunteers but of persuading people to let us collect confidential data from their computers in an intense data collection phase. We selected the organisations based on accessibility. However, we believe that the present sample represents knowledge work to the extent that is enough for conducting the study. The knowledge work tasks performed in these organisations may not represent all possible knowledge work tasks in the world. However, the types of tasks found in our data set are common to various types of knowledge work and thus probably generalisable beyond the limited convenience sample of participants and tasks.

All participants were volunteers. The participants who participated for more than a day were given two cinema tickets each to compensate for their time and effort. Feedback and preliminary results were given to all participants who were interested. For reasons beyond our control, only 1 day of data collection per person was possible in the city administration. Others were asked to participate for five working days but more were welcomed. Initially, the study had one more participant from a university, but her log data were corrupted and thus could not be analysed here.

### 3.2. Data collection and analysis phases

Several data collection methods were used, and all data types were collected from all participants. Logging, screen capture and direct observation happened during a continuous period of working days (participants' holidays excluded). The data collection phases were conducted in the following order:

1. Starting interview;
2. Daily questionnaires, logging and screen capture of task performance;
3. Direct observation (during one of the logging days);
4. Exit interview.

The methods and their use in later analysis are described in more detail below.

An *individual interview* was held with each participant before the data collection phase started. The purpose was to get background information of the participant and discuss her information behaviour. Interview data were not analysed per se but helped us understand the work of each participant.

The data on information resource use were collected in the participants' work places, while they were doing their normal, daily work tasks. At the beginning of each data collection day, the participants filled in an *electronic questionnaire* where they described their expected tasks of the day, their estimated knowledge of the task processes and estimated task complexity. In the analysis phase, this morning questionnaire was used to find out which work tasks might be identifiable in the log data. The questionnaire also included two out of three questions related to work task complexity. These two figures formed part of final work task complexity that was assigned to each work task by the researcher.

During the working days, the data were collected via *logging software* (ManicTime, http://www.manictime.com) and *screen capture software* (Snagit, https://www.techsmith.com/snagit.html) that were installed on the participants' computers. The logs collected include timestamps, the name of the information system used, a more precise name of the site, file or other location identifier and the duration of the stay. In the analysis phase, all dwell times were calculated from the transaction logs. Logs are rather plain, and rich screen capture videos were used to help connect each work task to corresponding rows in the log. Screen capture includes all that is visible on the screen (mouse clicks, texts and typing), so it clearly facilitates the process of understanding which work task is being performed and when. To correctly align work tasks to the events in the log is fundamental for successful analysis.

We also followed the work of each participant using *direct observation*, sometimes called shadowing [46]. Observation lasted for about one working day per each participant. Thus, the participants filled in questionnaires, and the logging and screen capture were run without interference of the researcher during most data collection days. The only exception was the municipal administration, because the only data collection day also included the direct observation. Maintaining the participants' trust and welfare was also a key point during direct observation; observation may make the participants feel uncomfortable. Thus, they were told that the researcher can leave earlier if needed, or she can

**Table 1.** Participants

| Organisation type (no. of organisations) | Org. description | No. of days (no. of participants, no. of females, no. of superiors) | Education/experience in work role[a] | Work roles |
|---|---|---|---|---|
| Commercial company (2) | Large retail and service chain/ small maintenance company | 37 (7, 6, 1) | Four Master's level degrees, two Bachelor's level degrees, one non-reported/work experience 1.5–3.5 years, one non-reported | Controller, HR business partner, concept manager, two category sourcing managers, CRM manager, development manager |
| University (3) | Multidisciplinary university/ technical university/university of applied sciences | 26 (4, 3, 0) | One doctor, three Master's level degrees/work experience 0.5–5 years | Senior lecturer, two doctoral students, researcher |
| City administration (1) | Large city in Finland | 10 (10, 7, 2) | Nine Master's level degrees, one non-reported/work experience 0.5–9.5 years | Three communications managers, specialised planner, two Web communications officers, senior HRD planning officer, communication and collaboration specialist, manager of human resource planning, skill development specialist |

HR: human resources; CRM: customer relationship management; HRD: human resource development.
[a]Often longer in the organisation or in similar tasks.

be absent for shorter times, for example, by not taking part in a meeting. During the observation day, the researcher was present in the office of the participant and took free-form notes of the participant's work. The key point of the observation was to learn to understand the work in practice since it was a prerequisite to correctly spot the work tasks in the flow of the log when analysing the data. The importance of this data collection method for later analysis cannot be overplayed. Following the work on the spot gives a deep understanding of what the participants' work is like in concrete terms, how often they switch between work tasks and so on, the participants can be asked about their work continuously, and they often talk about their work even without asking. Although all data collection days are not observed on the spot, the sample day helps interpret other days correctly.

At the end of each data collection day, the participants again filled in an *electronic questionnaire* similar to the one filled in the morning. They described the tasks that they had actually performed and estimated their complexity. This afternoon questionnaire provided information about which work tasks should at least be found in the data. Sometimes participants forgot to list some tasks that they had performed, so the researcher compared the morning and the afternoon questionnaires. It also happened that a planned task did not take place, or that the afternoon questionnaire revealed some unexpected tasks. Together, the two questionnaires included all three questions concerning task complexity. The researcher calculated an aggregate variable of them and assigned each task a task complexity value (see next section). Furthermore, both questionnaires included questions about information resource use which again helped to connect work tasks to corresponding rows in the logs (e.g. a specific information system is listed to be used only in one task); and in the afternoon questionnaire, the participants also gave us an estimate of the clock time at which they performed the task. It was an additional hint for us but of course only a suggestive one since the time may be hard to recall at the end of the day.

The data collection ended with an individual *exit interview* with each participant. The interview was held a few weeks after the logging had ended because the researcher needed to initially analyse the data before the interview. The interviews typically included discussion of where work task boundaries were if they were unclear to the researcher. Please note that these were not self-evident for the participants either; it is hard to remember the exact course of a day afterwards, even though assisted with information gained from the log and screen capture that were shown to the participants when necessary. Because in the city administration, the data collection lasted only for 1 day per participant and that day included direct observation, an exit interview was considered unnecessary.

The participants were given detailed explanations on how the whole study and the data collection software work. They were advised both orally and in writing how to control the logging and screen capture software; continuous possibility to support was provided by email and phone. Both logging and screen capture could be turned off and on, and the participants could delete rows and events from the log. They were advised to mark in the log through tags if they wanted the researcher to delete events from the log or the screen capture video. Or, they could otherwise inform the researcher about needs to modify the data. Needs for data modifications were inquired expressly once more when the data collection phase ended. Once the data were collected, the following analysis phases were performed:

1. Paring the work tasks in the morning and afternoon questionnaires;
2. Deciding which work tasks are visible in the log enabling further analysis and giving these work tasks a running identification number;
3. Calculating the complexity of each work task and assigning each work task to a work task type category;
4. Matching each work task to corresponding events in the log;
5. Listing all information resources found in the log and forming logical categories of them;
6. Calculating the length of use of each resource in each task;
7. Comparing dwell times between tasks.

Since the data were authentic with no explicit work task boundaries included, finding the boundaries of work tasks required a lot of intensive detective work (the analysis phase 4 in the above list) by the researcher. The data included 40,200 rows of log. Each row had to be linked to a corresponding work task. The various data types enabled this as explained earlier in this section. The key was to understand what the participants were doing in each row. Knowing this was made possible using the hints gained in the interviews, direct observation, screen capture and other rows of the log. Some rows were of course related to other things apart from the studied work tasks and thus excluded from further analysis. The next two sections provide detailed information about the analysis.

### 3.3. Defining task types and complexity

In this article, two independent task variables are of interest, namely, complexity and task type. To illustrate them, Appendix 1 includes a table with sample work tasks. *Task complexity* means here the participants' subjective experience and it was presented for the participants as such. It seemed rather easy and intuitive for them to estimate their tasks' subjective complexity. Sometimes they asked whether they should consider the tasks through somebody else's eyes. The instructions given were that they should estimate the complexity as themselves, as the authentic task doers. However, we also noted that if one considers one's own task familiar to oneself but highly complex for others, it may still increase the complexity estimate unintentionally.

In real life, there are numerous factors that can affect use of time in work task performance simultaneously. All of these cannot be measured exactly, given an authentic situation but (subjective) work task complexity, as shown to affect a range of information-related activities in earlier research (see the 'Literature' section), can act as a strong, aggregate/collective factor for them. In their subjective complexity estimates, the participants sum up the effects of several factors at once in the way they find relevant. After all, the task performers themselves are confronted with the actual tasks (c.f. Byström and Hansen [29]) and thus are able to estimate their complexity.

The task complexity was measured as the mean of three different task complexity measures on task questionnaire forms. Both in the morning and the afternoon forms, the participants reported how complex they felt each task will be and was. In the morning, they also assessed how much they knew about the task performance process. This figure was inverse to task complexity; the more they knew, the simpler the task. All estimates were given as percentages. If a figure was unavailable for some task, its complexity was calculated as the mean of the remaining estimates. If there was only one figure given (such as when a task had come up suddenly during the day), it was considered directly as the task's complexity. In all, 22% of potential estimates were missing in the final data set.

A continuous measure of task complexity is used when calculating correlation coefficients. However, in figures and tables, complexity has to be categorised. We used a simple categorisation that splits the tasks into four categories with about an equal number of tasks. The categories are I (tasks from 0% complexity to 21.7%, median 16.7), II (21.8%–38.3%, 30.0), III (38.4%–50%, 45.0) and IV (50.1%–100%, 60.0). This distribution shows that work tasks tend to be rather simple. The category IV has most variation in absolute work task complexity which can make the findings concerning complex tasks harder to interpret. However, these categories do not affect correlations. The categories only define how complex each task is compared with other tasks (e.g. the task belongs to the most simple quartile) which is a valid way of analysing the data. The skewness of the distribution is a natural feature of data, and we can expect that overall, very complex tasks are infrequent. Similar operationalisations of task complexity have been used in earlier literature (see section 'Work task complexity and type as empirical variables').

Besides task complexity, we formed four data-driven task type categories which can be applied to studies in other domains as well. The task type categorisation was formed ex post facto based on the task descriptions the participants had written in the electronic questionnaires. Initially, we considered features shared between the tasks and clustered tasks accordingly. Having formed stable clusters, we characterised their features and labelled them with descriptive titles, calling them task type categories. The categorisation is based on rather abstract and generic task features: (1) whether communication is involved, (2) whether the task is at the core of the substance of the work or rather supporting the main functions and (3) the degree of creativity or intellectual effort required. These task types also differ in the way information is used. Our categorisation does not include any domain-specific task types, thus being broadly applicable. Four task types were found.

*3.3.1. Communication tasks.* In communication tasks, communication is a dominant feature and a precondition for a successful task fulfilment, which is clearly stated in the task descriptions written by participants. Communication can be either one- or two-way. These include going through email and replying to the messages, or taking part in a meeting.

*3.3.2. Support tasks.* Support tasks support the core tasks of an organisation or of the participant in question. Routine-like, recurrent administrative tasks belong to support tasks as well as small tasks that by definition have a straightforward performance process. These include checking the number of participants to an event or taking care of monthly accounting. In support tasks, people do not (by definition) actually create new information (they may at most join existing information in a predefined manner) nor make important decisions. Note that support tasks are not automatically 'simple' by any means. For example, administrative tasks may prove complex because of lack of procedural knowledge, malfunctioning information systems, insufficient information available or the procedure may otherwise be much more complicated than the protocol (written or otherwise known or supposed) implies.

*3.3.3. Editing tasks.* In editing tasks, intellectual effort is needed when participants comment on the work done by others or edit something that has been initialised earlier. These include editing project forms or updating the organisation's statistics. Editing tasks are typically easily recognisable in the task descriptions due to the use of specific words (such as ‘commenting’, ‘finishing’, ‘editing’).

*3.3.4. Intellectual tasks.* Intellectual tasks are similar to editing tasks, but instead of editing, the participant indicates that something is created right from the start. Here creating means both immaterial and material outputs. Please note that typically work tasks are related to each other, and in practice, intellectual input builds on some earlier work. However, the way participants described this group of work tasks differed, for example, from editing tasks. The descriptions indicated a creative process or high intellectual input, not referring to earlier efforts. These include constructing monthly reports or calculating order and sales forecasts of a product in order to generate optimal sales campaigns.

In addition to these categories, there were two information seeking tasks and five tasks named ‘mixed assignments’ in the data. As these were so few, information seeking tasks were considered closest to intellectual tasks and mixed tasks were classified as support tasks. In the latter case, it is unlikely that demanding, larger tasks had been lumped together without an exact description.

The task types presented here are grounded in the data set used. However, they also have a theoretical connection to information use: one can assume that information resource use also differs between these task types. In communication tasks, the format of information and its processing are determined by the fact that its main function is to be communicated – often immediately – to other actors. In support tasks, information is often processed as an information object, that is, really instrumentally and not content-wise. In editing and intellectual tasks, information is processed more as a content matter. Especially in intellectual tasks, new information is created, often both in sense of knowledge and concrete output. Thus, support, editing and intellectual tasks form a continuum of increasing creativeness.

Since the type categories were assigned to tasks based on case-specific consideration, both between-classifiers and within-classifier reliability tests were conducted. Two other researchers from the same research centre got instructions and descriptions of work task categorisation similar to the description presented earlier in this article. First, they got five random tasks to train with and feedback on success. Then, they reclassified a random sample of 50 tasks (Cohen’s Kappa: 0.69); the primary classifier reclassified 50 random tasks (Cohen’s Kappa: 0.83), as well. These indicate satisfactory correspondence. Our earlier paper [7] discussed the fact that work task complexity and work task type are connected in an interesting way. Many support tasks are simple, many editing tasks are semi-simple and intellectual tasks are often complex. Many communication tasks are semi-complex.

### *3.4. Measuring resource use*

In this article, we operationalise information resource use as dwell time visible in the log. Thus, dwell time is a measure of how many seconds a participant kept a resource open as an active window during a work task. Rather than relying solely on the participants’ narratives of what they did, we collected a log that revealed what they in fact did, since in practice, work task performance consists of concrete time dwelt in various information resources. The analyses conducted here do not reveal exact actions of participants (e.g. reading or writing), but certainly the resource categories applied give us a preliminary understanding of how the time was spent when conducting the work tasks. This is something that cannot be exactly analysed using self-reports; if we wish to know concrete task performance, dwell time measured through logging is a good indicator.

The logging software collected information about the information systems used and the time expended on each system. We classified the information systems into information resource categories. In developing the classification, first, all the resources in the data were listed and then the resource categories were formed using the following principles:

- Every information resource has a suitable category.
- All resource categories are represented in the data.
- The categories do not overlap.
- The use of each resource category has the potential to react to task complexity and type.
- Each category forms a qualitatively sensible group of similar information resources based on their surface-level functions, the intended use, the contents and access to the information available.
- The resource categories are similar enough to other studies to make comparisons.
- The categories have enough total dwell time to be compared with other categories.
- The categories reflect the range of information resources of the participants.
- They respond to our research questions and interests.
- They are generic enough to suit all studied organisations.

Although our categorisation is data-driven, it is not specific to our study but similar to the ones used in earlier literature as well (see section 'Information resources') and applicable to future studies. Our five main categories for information resources are as follows:

- Communication;
- Web search engines;
- Other websites;
- PC;
- Organisational resources.

Communication resources include email and instant messaging software. Web search engines are found in the public, free Web. The dwell time in search engines includes only the actual time spent writing a query and browsing the results. This time does not include following any links outside. Please note that typically many types of websites, such as a company's homepage, include an internal search feature, meant to help in navigating, but that does not make them 'search engines' in our categorisation. Search engines are websites meant for finding textual or other types of data in the whole Web or in specialised areas. Organisational resources include various databases, special software purchased by the organisation and the intranet. PC includes basic office tools and freeware on participants' computers, and the settings and directory structure of the computer.

The classification is intended to be applicable over all studied organisations as well as easily comparable with results from other studies. Thus, it cannot include any substance-specific categories. Furthermore, the relation of substance-specific categories with task complexity is questionable: for example, a spreadsheet is probably to be used for calculating and a word processor for writing independent of task complexity. These actions are inherently more related to the contents of the task which is not of interest in this article. The same concerns organisational resources. In one organisation, accounting software is needed and survey software in another. Perhaps it would have been possible to categorise organisational databases more specifically. However, this was considered not feasible for the above-mentioned reasons and because the authors are not experts in the domains studied. We neither had access to these databases to support a more specific categorisation.

## 4. Results

This section includes the detailed analyses whereas section 'Discussion' discusses the findings on a more general level. 'Results' section is organised as follows. First, we present the effects of task complexity and task type on information resource use. Then, we present the effects of task complexity within work task types. Tables in this section present (a) average task duration, (b) average dwell time of each resource in seconds per task and (c) average percent of task duration. Please note that (c) is not calculated as (b)'s proportion of (a). Instead, (c) is *micro average* meaning that the percentages are first calculated separately in each task and then averaged. Thus, each task has an equal weight. An interested reader can easily calculate the macro average as (b)'s proportion of (a). Macro averages are not presented here because, problematically, long tasks weigh more in them than short tasks. We use Pearson's correlation coefficient to calculate the linear connection between task complexity and resource use.

### 4.1. The effects of task complexity on information resource use

Table 2 shows the average lengths of time dwelt in each information resource across task complexity categories (I–IV). Table 2 also shows the average share of task duration each resource type has. First, the more complex the task, the longer it lasts. This seems quite self-evident; complex tasks cannot be performed quickly. The use of PC resources increases and the share of organisational resources decreases with growing task complexity. The use of communication resources is connected to work task complexity but not in a linear fashion: communication resources are used clearly more in work tasks of middle complexity than in simple or complex tasks. The use of search engines increases with growing task complexity but the share of total task duration is marginal.

### 4.2. The effects of task complexity within task types

In this section, the effects of task complexity are separately analysed within task types, namely, communication, support, editing and intellectual tasks. First, we give an overview of resource use in varying task types.

**Table 2.** Average dwell time of information resources across task complexity categories

| | Information resources | Task complexity | | | | Correlation |
|---|---|---|---|---|---|---|
| | | I | II | III | IV | |
| Average task duration (s) | | 2885 | 3662 | 4671 | 5431 | 0.21** |
| Average use (s) per task | Communication | 699 | 1280 | 1724 | 975 | 0.05 |
| | Search engines | 8 | 19 | 43 | 73 | 0.14* |
| | Other Web | 410 | 278 | 134 | 351 | – 0.02 |
| | Organisation | 767 | 532 | 617 | 540 | – 0.03 |
| | PC | 1001 | 1555 | 2152 | 3493 | 0.27** |
| Average % of task duration | Communication | 27.2 | 38.4 | 37.9 | 26.9 | – 0.05 |
| | Search engines | 0.3 | 0.5 | 0.7 | 0.8 | 0.12 |
| | Other Web | 10.0 | 8.2 | 5.2 | 5.6 | – 0.08 |
| | Organisation | 26.0 | 15.8 | 15.0 | 13.8 | – 0.12 |
| | PC | 36.5 | 37.0 | 41.2 | 52.8 | 0.15* |
| | n (work tasks) | 54 | 65 | 80 | 72 | N 271 |

PC: personal computer.
Pearson's correlation is calculated using continuous task complexity measure.
$*p \leq 0.05$; $**p \leq 0.01$.

**Table 3.** Average dwell time of information resources by task type

| | Information resources | Task type | | | | The differences between classes |
|---|---|---|---|---|---|---|
| | | Communication | Support | Editing | Intellectual | |
| Average task duration (s) | | 3691 | 3393 | 4323 | 5343 | K-W |
| Average use (s) per task | Communication | 2179 | 723 | 828 | 736 | K-W** |
| | Search engines | 7 | 26 | 16 | 96 | K-W |
| | Other Web | 164 | 245 | 409 | 320 | K-W |
| | Organisation | 370 | 803 | 798 | 600 | K-W |
| | PC | 971 | 1596 | 2271 | 3590 | K-W** |
| Average % of task duration | Communication | 57.6 | 24.0 | 25.1 | 17.3 | K-W** |
| | Search engines | 0.4 | 0.4 | 0.3 | 1.3 | K-W |
| | Other Web | 7.8 | 7.1 | 9.1 | 4.3 | K-W |
| | Organisation | 10.6 | 22.5 | 19.3 | 19.5 | K-W |
| | PC | 23.7 | 45.9 | 46.3 | 57.7 | B-F** |
| N 271 | n (work tasks) | 86 | 41 | 66 | 78 | |

PC: personal computer; K-W = Kruskal–Wallis test; B-F = Brown–Forsythe test.
$*p \leq 0.05$; $**p \leq 0.01$.

The average dwell times in information resources across task types are shown in Table 3. Intellectual tasks last longest. The differences between task types are not statistically significant but between communication and intellectual, and support and intellectual tasks they are suggestive. The time spent in PC resources is smallest in communication tasks and grows from support to editing and intellectual tasks. Pairwise comparisons suggest that in terms of *relative use*, in communication tasks, PC is used less than in other tasks; intellectual tasks differ from editing tasks and almost statistically significantly from support tasks, while support and editing tasks do not differ from each other. The pairwise comparisons of PC use in *seconds* are similar but suggest weaker differences overall: communication tasks differ significantly from editing and intellectual tasks and almost significantly from support tasks. Editing and intellectual tasks differ almost significantly from each other, whereas there is no connection between editing and support, or support and intellectual tasks. Overall, PC is the most common information resource in other tasks than communication tasks, where communication resources are used most.

**Table 4.** Average dwell time of information resources in communication tasks

| | Information resources | Task complexity | | | | Correlation |
|---|---|---|---|---|---|---|
| | | I | II | III | IV | |
| Average task duration (s) | | 2938 | 2722 | 4948 | 3356 | 0.23* |
| Average use (s) per task | Communication | 1467 | 1861 | 2884 | 1898 | 0.19 |
| | Search engines | 14 | 6 | 7 | 4 | − 0.08 |
| | Other Web | 150 | 161 | 179 | 150 | 0.01 |
| | Organisation | 500 | 204 | 399 | 460 | 0.09 |
| | PC | 806 | 489 | 1479 | 844 | 0.16 |
| Average % of task duration | Communication | 53.2 | 61.6 | 56.2 | 58.3 | 0.00 |
| | Search engines | 0.3 | 0.7 | 0.2 | 0.1 | − 0.11 |
| | Other Web | 5.0 | 11.6 | 7.9 | 3.9 | − 0.06 |
| | Organisation | 18.2 | 6.5 | 8.3 | 14.8 | 0.02 |
| | PC | 23.3 | 19.6 | 27.5 | 22.8 | 0.03 |
| | **n (work tasks)** | **15** | **25** | **32** | **14** | **N 86** |

PC: personal computer.
Pearson's correlation is calculated using continuous task complexity measure.*
p ≤ 0.05; **p ≤ 0.01.

**Table 5.** Average dwell time of information resources in support tasks

| | Information resources | Task complexity | | | | Correlation |
|---|---|---|---|---|---|---|
| | | I | II | III | IV | |
| Average task duration (s) | | 2309 | 3834 | 4323 | 3024 | 0.10 |
| Average use (s) per task | Communication | 463 | 463 | 1055 | 975 | 0.29 |
| | Search engines | 5 | 49 | 43 | 2 | 0.00 |
| | Other Web | 95 | 521 | 132 | 301 | 0.01 |
| | Organisation | 750 | 968 | 489 | 1194 | 0.02 |
| | PC | 997 | 1834 | 2604 | 552 | − 0.02 |
| Average % of task duration | Communication | 17.0 | 14.7 | 28.8 | 41.4 | 0.42** |
| | Search engines | 0.1 | 0.6 | 0.8 | 0.1 | 0.03 |
| | Other Web | 4.3 | 7.1 | 7.8 | 10.7 | 0.07 |
| | Organisation | 32.8 | 19.7 | 10.5 | 29.4 | − 0.17 |
| | PC | 45.7 | 57.9 | 52.2 | 18.4 | − 0.24 |
| | **n (work tasks)** | **12** | **10** | **12** | **7** | **N 41** |

PC: personal computer.
Pearson's correlation is calculated using continuous task complexity measure.
*p ≤ 0.05; **p ≤ 0.01.

*4.2.1. Communication tasks.* Table 4 summarises the resource use in communication tasks. Task duration is positively correlated with task complexity in communication tasks. However, task complexity does not have a linear effect on the length of use of any information resources. As expected, the share of communication resources is not affected by task complexity since communication is an integral part of these tasks by definition. The absolute time (seconds) using communication resources and PC seems to vary by task complexity. However, this is hard to interpret since average task duration changes at the same time in a somewhat similar fashion. This finding may be interpreted so that in semi-complex tasks (category III), the increased task duration is caused by pronounced use of communication and PC resources.

It may be that task complexity's effects on resource use in communication tasks are so weak because this group of tasks is quite heterogeneous though communication tasks form a clearly distinguishable work task category in participants' descriptions. Communication tasks may vary from routine email checking to instructing other people and from meetings to replying to a complicated email inquiry that takes hours to complete.

*4.2.2. Support tasks.* Table 5 summarises the resource use in support tasks. The more complex a support task, the more interaction with other people is needed, that is, the use of communication resources increases. Email enables repository-

**Table 6.** Average dwell time of information resources in editing tasks

| | Information resources | Task complexity | | | | Correlation |
|---|---|---|---|---|---|---|
| | | I | II | III | IV | |
| Average task duration (s) | | 3563 | 4873 | 4561 | 4182 | 0.10 |
| Average use (s) per task | Communication | 418 | 1272 | 1052 | 463 | − 0.04 |
| | Search engines | 8 | 16 | 17 | 26 | 0.24 |
| | Other Web | 1080 | 384 | 84 | 72 | − 0.19 |
| | Organisation | 809 | 712 | 1163 | 509 | 0.02 |
| | PC | 1249 | 2489 | 2245 | 3113 | 0.21 |
| Average % of task duration | Communication | 15.7 | 30.0 | 33.3 | 20.0 | − 0.02 |
| | Search engines | 0.4 | 0.4 | 0.2 | 0.2 | − 0.00 |
| | Other Web | 23.4 | 7.6 | 0.9 | 4.5 | − 0.20 |
| | Organisation | 19.7 | 22.3 | 19.2 | 15.2 | − 0.01 |
| | PC | 40.9 | 39.7 | 46.5 | 60.1 | 0.15 |
| | **n (work tasks)** | **16** | **19** | **16** | **15** | **N 66** |

PC: personal computer.
Pearson's correlation is calculated using continuous task complexity measure.
$^{*}p \leq 0.05$; $^{**}p \leq 0.01$.

like use, as well, but instant messaging is mainly used for direct communication. Although support tasks are often considered as routine, if any problems occur or if the task is unfamiliar to the task performer, other people are consulted.

When support tasks become more complex, PC resources are used proportionally less ($r = -0.24$; $p > 0.05$). PC use is frequent but drops clearly in the most complex support tasks. Perhaps consulting colleagues and using organisational information systems replace files that can be accessed locally.

*4.2.3. Editing tasks.* Task complexity is not a strong indicator of resource use in editing tasks (see Table 6). It seems that editing tasks are a heterogeneous group like communication tasks. Common to all editing tasks is the need to modify or comment on some information object, but unlike in intellectual tasks, the quantity of work needed may vary in editing tasks from cursory reading to holistic restructuring. In addition, the effort and the resources needed may be different depending on whether or not the initial creator is the same person as the editor. The absolute time spent in search engines slightly increases with growing task complexity. This is especially interesting because the time spent on other websites decreases. Also, the use of PC resources increases with an increase in the task complexity. However, no correlations are statistically significant.

*4.2.4. Intellectual tasks.* Task complexity affects the performance of intellectual tasks (see Table 7). Complex intellectual tasks take much longer to perform than simple ones. This is mainly caused by the increased use of PC resources. The finding is understandable since PC resources are typically used for writing and computing, that is, creating information. The relative use of organisational resources decreases with an increase in the task complexity. It may indicate that basic PC tools are more important in the creative process, whereas organisation's more special resources suit other tasks better.

It is interesting what happens to communication when intellectual tasks increase in complexity. Seemingly, the time used in communication increases (comparing averages I–IV), but because the increase in overall task duration is so clear, the share of communication actually drops moderately. Perhaps since intellectual tasks often encompass a creative process, this process, especially if considered complex, requires contemplation which seldom can be assisted by messages to colleagues. In the intellectual tasks of our data, often the task performer herself seemed to be the best expert in the subject.

## 5. Discussion

### *5.1. Task complexity*

*Task complexity* is related to the use of some information resources. In the simplest tasks, most time is spent rather evenly in communication, organisational and PC resources. In middle complexity categories, organisational resources drop in use and they seem to be replaced by communicational resources. In complex tasks, the dwell time in

**Table 7.** Average dwell time of information resources in intellectual tasks

| | Information resources | Task complexity | | | | Correlation |
|---|---|---|---|---|---|---|
| | | I | II | III | IV | |
| Average task duration (s) | | 2453 | 3553 | 4525 | 7227 | 0.25* |
| Average use (s) per task | Communication | 321 | 716 | 809 | 829 | 0.02 |
| | Search engines | 2 | 25 | 121 | 133 | 0.16 |
| | Other Web | 132 | 137 | 103 | 555 | 0.12 |
| | Organisation | 1089 | 567 | 607 | 456 | − 0.18 |
| | PC | 910 | 2107 | 2885 | 5254 | 0.33** |
| Average % of task duration | Communication | 19.9 | 21.9 | 17.9 | 14.8 | − 0.17 |
| | Search engines | 0.2 | 0.4 | 1.8 | 1.5 | 0.20 |
| | Other Web | 3.7 | 2.5 | 2.9 | 5.7 | 0.09 |
| | Organisation | 38.3 | 22.3 | 25.0 | 9.8 | − 0.28* |
| | PC | 37.9 | 52.9 | 52.5 | 68.2 | 0.31** |
| | **n (work tasks)** | **11** | **11** | **20** | **36** | **N 78** |

PC: personal computer.
Pearson's correlation is calculated using continuous task complexity measure.
*$p \leq 0.05$; **$p \leq 0.01$.

communication resources seems to be superseded by dwell time in PC resources. These findings indicate that in simple tasks, all resources are equally important and information interaction is versatile. In more complex tasks, this diversity changes towards more interaction between people, whereas in the most complex tasks, task performers need to concentrate on (solitary) production of information.

Task duration grows steadily with task complexity and almost doubles from simple to complex tasks. This is especially caused by the increased use of PC resources. This is contrary to some earlier findings [38] that did not suggest a connection between task complexity and *frequency* of PC use; Saastamoinen et al. [38] actually suggested that PC's *importance* drops with growing task complexity. Partly this may be explained by differences in data collection and analysis methods. However, the work tasks studied also differed. The present data represent a rather heterogeneous set of tasks, whereas [38] covered a small number of planning and administrative tasks in one organisation.

In the present data, the dwell time in searching the Web increases a little with task complexity, though overall the dwell time in search engines is quite small. Searching does not consume the time but its consequences, such as exploiting the information found when writing on PC may do so. In real-life tasks, in contrary to IIR tests, the task performers must finish the task a way or another. Thus, if searching is successful, the dwell time in Web search engines is probably to be small. If it is not successful, the participant is not probably to continue searching for a long time but tries another resource, unless the Web is the best resource she can come up with in that situation.

Interestingly, communication resources were used the least in the most simple and most complex tasks. Earlier, we [38] found that task complexity increases slightly the use of communication resources. In the present data, our findings suggest this happening only in support tasks where the connection is quite clear. This finding is interesting because the data set in Saastamoinen et al.'s [38] work was collected in a city administration context which might indicate that the majority of work tasks studied were actually support tasks. Furthermore, the study [38] was based on the frequency of use rather than the duration of use, and the category of communication included face-to-face meetings because the data collection methods were different.

Also, Byström's [3] findings showed that increasing task complexity increases clearly the use of people as information sources. However, the comparison of these findings is not straightforward. Byström's [3] data were collected 20 years ago in a single organisation (municipal administration) which probably affects the findings. Furthermore, she analysed the frequency of use of each resource, not the time spent using them. Our category of communication includes the use of email and instant messaging, whereas Byström's [3] category is rather based on face-to-face communication which is not visible in the present data set.

## *5.2. Task types*

Profession-independent *work task types* are seldom used as an independent variable in empirical information seeking studies. We found that work task type is connected to information resource use. Task duration increases from support to editing and intellectual tasks; these task types in this order require progressively more intellectual input by definition, as explained in section 'Defining task types and complexity'. The increase in total task performance time is mainly caused by the increased use of PC resources. In communication tasks, communication software is used over twice as much as in other task types, that is, nearly 60% of task duration.

Communication resources often belong to the most used ones [43]. In our data, based on time spent in communication software, this finding holds best among communication tasks. This may be due to the differences in methods: communication may be an important source of information and thus easily named in surveys and interviews; of course our data do not include face-to-face communication or via telephone. However, it still seems that the time spent in communication software does not totally reflect the often reckoned importance of communication as a means of information seeking. Dwell time and effectiveness (of information interaction) may not always correlate.

### *5.3. Task complexity within task types*

Task complexity affects information resource use differently within task types. Perhaps the most interesting finding was that in intellectual tasks, the effects of task complexity are clearer than in other tasks: task duration and the use of PC increase substantially with task complexity and the use of organisational resources decreases. This may indicate that the performance process of intellectual tasks is more open than in other task types, which leaves leeway for the effects of task complexity. Another explanation is that intellectual tasks represent a homogeneous set of tasks, and thus the effects of task complexity are easily visible.

On the contrary, communication tasks and editing tasks do not react much to varying task complexity in terms of statistical significance. In support tasks, complexity increases the use of communication software. This may indicate that the performance process is known in theory, but as the complexity increases, some unexpected problems may occur and consulting other people is needed. For example, Kumpulainen and Järvelin [47] analysed various kinds of barriers in task performance and found that complex tasks include more barriers than simple ones, and these often relate to the use and integration of information systems.

In editing tasks, some trends are visible though the correlations are not statistically significant. Especially search engines, in terms of absolute time, are used more but other websites less when task complexity increases. This may indicate increase in unsuccessful searching since search engine use takes a larger proportion of all Web use.

## 6. Limitations and implications

This study has some limitations especially due to the data collection methods. The data set collected could always be bigger; as this was a real-life study, it is possible that other context factors (e.g. the organisation type) affected information resource use besides task complexity and task type. In an ideal situation, the data would have been larger, covering several different organisations, and these organisation types could have been compared. Due to practical reasons, this was not feasible. Another solution could have been to restrict the data collection into one organisation type, such as only analysing the data from the two commercial companies, which is a common approach. Perhaps the findings would have been more consistent. However, settling on a single (type of) organisation could not have assisted in reaching the goal of the study: to understand information resource use in knowledge work. We did not have an initial reason to expect that, considering the phenomena and variables analysed, the organisations chosen should be very different. Furthermore, the features of organisations that could affect information resource use were not known. The organisations and participants studied differed in some respects but also had much in common. We analysed task and resource features that should be applicable to several types of knowledge work.

Our analysis comprised only the use of computer-based information resources which is a major limitation. Direct observation showed that papers and printed books are sometimes used in task performance, but the same information is often found in electronic form as well. However, the decision to analyse only computer-based resources most certainly decreased the amount of communication in the data since all telephone calls and face-to-face meetings were missed.

We did not analyse the type of documents or information (internal reports, news headlines and contact information) that the participants interacted with using various resources. We neither analysed various facets of information interaction (e.g. reading, synthesising and reporting [6]) that summed up to the measured dwell times. These data for such analyses are available especially in the recorded screen videos, but their analysis is left for later study.

Despite these limitations, we contribute findings that are based on real-life information resource use rather than its narratives. We also took a step towards understanding the relationship between other information seeking and Web information retrieval since we studied Web search engines as one information resource among others. The data we have reported indicate the dwell times and their distributions but not the number of use instances of each resource that summed up to the total times. We leave for later study to analyse the frequencies of resource uses.

As we have shown, the time spent in various information resources varies by task type and complexity. Organisations could consider whether the time spent in various resources is effective in view of the tasks at hand. Our study suggests that long times are overall spent in PC resources and in communication resources, and one could, for example, elaborate: (1) whether the time in PC resources is really productive, or a consequence of evading barriers [47], such as a poor information system; (2) whether the time spent in communication resources is communicating the core aspects of the work or such as answering to secondary email inquiries that could or should be arranged in a different way; (3) whether organisational information systems are actually underused, that is, there exists a system with relevant information but people find it too hard to use or do not know that it exists. This type of elaboration of time expenditure can help develop better information environments.

## 7. Conclusion

This article presented results of a field study analysing how work task complexity and work task type are connected to information resource use. By exploiting client-side log analysis, this study went beyond typical information seeking studies that analyse only narratives of information resource use, that is, questionnaires and interviews. We used logs to measure exact dwell times on all computer-based resources, including search engines. This information was united with contextual knowledge about work tasks. This exemplifies a simply applicable way to study task-based information seeking or interaction including data collection and analysis methods.

We were able to show that task complexity is a valuable independent variable in analysing information resource use but it is accompanied by other factors, as well. The proposed work task type categorisation sheds more light on the effects of work task complexity. The relationships between the use of various information resources and these work task features proved complex indeed. In particular,

- In terms of dwell times, search engines are used quite little in daily work tasks, but the use is pronounced in intellectual and complex work tasks.
- The average task duration increases from support tasks to editing and intellectual tasks and this is mainly due to the increased use of PC resources.
- In intellectual tasks, the effects of task complexity are clearer than in other tasks; among others, complex intellectual tasks involve more searching and PC use while less use of organisational resources.
- The dwell time in communication resources clearly increases with task complexity only in support tasks.

Future research could deepen the findings. For this, we suggest that a larger data set be collected in order to statistically analyse the explanatory power of each variable. We further propose the analysis of time management. It is crucial to understand whether the dwell time in a resource is beneficial or ineffective and where the biggest improvements in terms of effectiveness and efficiency of task performance may be achieved. The screen video also allows us to study *what* was actually done with each resource. It may be beneficial to extend the scope of studies from mere information seeking or searching towards understanding work tasks as a process with other phases, as well [15]. Studies like the present one provide a realistic view on information interaction in work task performance for use in both theory formation and developing better information infrastructures for knowledge work.

### Acknowledgements

The authors thank Ms Heljä Franssila (University of Tampere) for her cooperation in data collection.

### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

### References

[1] Case DO. *Looking for information: a survey of research on information seeking, needs, and behavior*. 1st ed. Amsterdam: Academic Press, 2002.

[2] Li Y and Belkin NJ. A faceted approach to conceptualizing tasks in information seeking. *Inform Process Manag* 2008; 44: 1822–1837.
[3] Byström K. *Task complexity, information types and information sources*. PhD Thesis, University of Tampere, Tampere, 1999.
[4] Liu J, Kim CS and Creel C. Exploring search task difficulty reasons in different task types and user knowledge groups. *Inform Process Manag* 2015; 5: 273–285.
[5] Wildemuth B, Freund L and Toms EG. Untangling search task complexity and difficulty in the context of interactive information retrieval studies. *J Doc* 2014; 70: 1118–1140.
[6] Järvelin K, Vakkari P, Arvola P et al. Task-based information interaction evaluation: the viewpoint of program theory. *ACM T Inform Syst* 2015; 33(1): article no. 3.
[7] Saastamoinen M and Järvelin K. Search task features in work tasks of varying types and complexity. *J Assoc Inf Sci Technol*, in press.
[8] Saastamoinen M and Järvelin K. Queries in authentic work tasks: the effects of task type and complexity. *J Doc* 2016; 72: 1052–1077.
[9] Kuhlthau CC. Towards collaboration between information seeking and information retrieval (paper 225). *Inform Res* 2005; 10.
[10] Järvelin K and Ingwersen P. Information seeking research needs extension towards tasks and technology (paper 212). *Inform Res* 2004; 10.
[11] Vakkari P. Task complexity, information types, search strategies and relevance: integrating studies on information seeking and retrieval. In: Wilson TD and Allen DK (eds) *Exploring the contexts of information behaviour*. London: Taylor Graham Publishing, 1999, pp. 35–54.
[12] Martzoukou K. A review of web information seeking research: considerations of method and foci of interest (paper 215). *Inform Res* 2005; 10.
[13] Broder A. A taxonomy of web search. *ACM Sigir Forum* 2002; 36: 3–10. New York, NY: ACM.
[14] Li Y. Exploring the relationships between work task and search task in information search. *J Am Soc Inform Sci Tech* 2009; 60: 275–291.
[15] Bartlett JC and Toms EG. Developing a protocol for bioinformatics analysis: an integrated information behavior and task analysis approach. *J Am Soc Inform Sci Tech* 2005; 56: 469–482.
[16] Du JT. The information journey of marketing professionals: incorporating work task-driven information seeking, information judgments, information use, and information sharing. *J Assoc Inf Sci Technol* 2014; 65: 1850–1859.
[17] Kallehauge J. Stage-driven information seeking process: value and uncertainty of work tasks from initiation to resolution. *J Inf Sci* 2010; 36: 242–262.
[18] Ibrahim AE. Use and user perception of electronic resources in the United Arab Emirates University (UAEU). *Libri* 2004; 54: 18–29.
[19] Claypool M, Le P, Waseda M et al. Implicit interest indicators. In: *Proceedings of ACM intelligent user interfaces conferences (UIU'01)*, Santa Fe, NM, 14–17 January 2001, pp. 33–40. New York: ACM.
[20] Kim J, Oard DW and Romanik K. User modeling for information access based on implicit feedback. In: *Proceedings of ISKO-France July 5–6 2001*, Nanterre: Université de Paris X, 2001, pp. 1–11.
[21] Konstan J, Miller B, Maltz M et al. GroupLens: applying collaborative filtering to Usenet news. *Commun* ACM 1997; 40: 77–87.
[22] Kellar M, Watters C, Duffy J et al. Effect of task spent reading as an implicit measure of interest. *Proc Am Soc Inf Sci Technol* 2004; 41: 168–175.
[23] Kelly D and Belkin NJ. Reading time, scrolling and interaction: exploring implicit sources of user preference for relevance feedback. In: *Proceedings of the 24th annual international ACM conference on research and development in information retrieval (SIGIR'01)*, New Orleans, LA, 1 September 2001, pp. 408–409. New York: ACM.
[24] Borlund P, Dreier S and Byström K. What does time spent on searching indicate? In: *Proceedings of the 4th information interaction in context symposium*, Nijmegen, 21–24 August 2012, pp. 184–193. New York: ACM.
[25] Liu J. User assessment of search task difficulty: relationships between reasons and ratings. *Libr Inform Sci Res* 2015; 37: 329–337.
[26] Maynard DC and Hakel MD. Effects of objective and subjective task complexity on performance. *Hum Perform* 1997; 10: 303–330.
[27] Courtright C. Context in information behavior research. *Annu Rev Inform Sci* 2007; 41: 273–306.
[28] Vakkari P. Task-based information searching. *Annu Rev Inform Sci* 2003; 37: 413–464.
[29] Byström K and Hansen P. Work tasks as units for analysis in information seeking and retrieval studies. In: *Proceedings of the 4th international conference on conceptions of library & information science*, Seattle, WA, 21–25 July 2002, pp. 239–251. Greenwood, CO: Libraries Unlimited.
[30] Kim S and Soergel D. Selecting and measuring task characteristics as independent variables. *Proc Am Soc Inf Sci Technol* 2005; 42.
[31] Byström K and Hansen P. Conceptual framework for tasks in information studies. *J Am Soc Inform Sci Tech* 2005; 56: 1050–1061.

[32] Li Y and Belkin NJ. An exploration of the relationships between work task and interactive information search behavior. *J Am Soc Inform Sci Tech* 2010; 61: 1771–1789.
[33] Pharo N. A new model of information behaviour based on the search situation transition schema (paper 203). *Inform Res* 2004; 10.
[34] Taylor A. Examination of work task and criteria choices for the relevance judgment process. *J Doc* 2013; 69: 523–544.
[35] Toms EG, O'Brien H, Mackenzie T et al. Task effects on interactive search: the query factor. In: *Proceedings of the 6th international workshop of the initiative for the evaluation of XML retrieval (INEX 2007)*, Dagstuhl Castle, 17–19 December 2008, pp. 359–372. Berlin: Springer.
[36] Liu P and Li Z. Task complexity: a review and conceptualization framework. *Int J Ind Ergonom* 2012; 42: 553–568.
[37] Kumpulainen S and Järvelin K. Information interaction in molecular medicine: integrated use of multiple channels. In: *Proceedings of the third symposium on information interaction in context*, New Brunswick, NJ, 18–21 August 2010, pp. 95–104. New York: ACM.
[38] Saastamoinen M, Kumpulainen S and Järvelin K. Task complexity and information searching in administrative tasks revisited. In: *Proceedings of the 4th information interaction in context symposium*, Nijmegen, 21–24 August 2012, pp. 204–213. New York: ACM.
[39] Du JT, Liu YH, Zhu QH et al. Modelling marketing professionals' information behaviour in the workplace: towards a holistic understanding (paper 560). *Inform Res* 2013; 18.
[40] Hansen P. *Task-based information seeking and retrieval in the patent domain: processes and relationships*. PhD Thesis, University of Tampere, Tampere, 2011.
[41] Saastamoinen M and Kumpulainen S. Expected and materialised information source use by municipal officials: intertwining with task complexity (paper 646). *Inform Res* 2014; 19.
[42] Savolainen R. Source preferences in the context of seeking problem-specific information. *Inform Process Manag* 2008; 44: 274–293.
[43] Hertzum M. Expertise seeking: a review. *Inform Process Manag* 2014; 50: 775–795.
[44] Nicholas D, Williams P, Rowlands I et al. Researchers' e-journal use and information seeking behaviour. *J Inf Sci* 2010; 36: 494–516.
[45] Taylor RS. Question-negotiation and information seeking in libraries. *Coll Res Libr* 1968; 29: 178–194.
[46] Czarniawska-Joerges B. *Shadowing: and other techniques for doing fieldwork in modern societies*. 1st ed. Copenhagen: Copenhagen Business School Press, 2007.
[47] Kumpulainen S and Järvelin K. Barriers to task-based information access in molecular medicine. *J Am Soc Inform Sci Tech* 2012; 63: 86–97.

## Appendix 1

**Table 8.** Sample work tasks

| Organisation | Work task label(s) by participant (morning/afternoon)[a] | Task type category | Complexity calculation (mean of morning complexity, inverse process knowledge and afternoon complexity) | Further details |
|---|---|---|---|---|
| Company | 'Reading and taking care of emails (my own and of the customer relationship management support)'/'emails' | Communication | $(50 + 1)/2 = 25.5 \Rightarrow$ complexity category II | The participant did not indicate the process knowledge. |
| Company | 'Checking the situation of an invitation in an information system' | Support | $1/1 = 1.0 \Rightarrow$ I | The task was not expected in the morning. |
| Company | 'Issue with marketing tool' | Editing | $50/1 = 50.0 \Rightarrow$ III | An unfinished job that has been started earlier; not expected in the morning. |
| University | 'Editing a paper draft'/'editing a paper' | Editing | $(100 + 50 + 100)/3 = 83.3 \Rightarrow$ IV | |
| University | 'Editing a paper draft'/'editing and writing a paper' | Intellectual | $(100 + 50 + 100)/3 = 83.3 \Rightarrow$ IV | The task turned out to be about writing new text as well as making revisions. |
| University | 'Teaching'/'teaching' | Communication | $(50 + 20 + 50)/3 = 40.0 \Rightarrow$ III | A part of teaching task happened via computer and could thus be logged. |
| City | 'Annulling the media monitoring contract'/'canceling the media monitoring contract' | Support | $(90 + 70 + 50)/3 = 70.0 \Rightarrow$ IV | |
| City | 'An article about rescue exercise for the intranet, retrieving pictures and making the article'/'news about rescue exercise' | Intellectual | $(50 + 30 + 50)/3 = 43.3 \Rightarrow$ III | |

[a]Work task labels are direct translations of authentic task descriptions given by the participants.