



HAL
open science

Influence of task context and task complexity on students' task model quality, search behavior and search outcome

Delphine Oger, Mylène Sanchiz

► To cite this version:

Delphine Oger, Mylène Sanchiz. Influence of task context and task complexity on students' task model quality, search behavior and search outcome. Reading and Writing, 2025, 10.1007/s11145-025-10639-7. hal-04951523

HAL Id: hal-04951523

<https://hal.science/hal-04951523v1>

Submitted on 17 Feb 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Influence of task context and task complexity on students' task model quality, search behavior and search outcome

Delphine Oger & Mylène Sanchiz

Laboratoire CeRCA (UMR 7295), Université de Poitiers

Version finale acceptée du manuscrit, pour dépôt dans HAL

Abstract

Searching for information with a search engine requires students to develop a coherent representation of their search goal. According to the RESOLV model of purposeful reading, students can draw on clues from the task context and the task instructions. This study examined the influence of task context and task complexity on students' initial task model, and the interplay between task context and students' task models on search behavior and outcomes when using a search engine. Ninety university students searched for information for two simple and two complex search tasks, conducted in both academic and non-academic contexts. **Students' initial task models** were mostly composed of core topic elements and of the surface-level actions required to search for information. Planning, regulation and standards of coherence were less frequently included in students' initial task models. For simple tasks, higher standards of coherence promoted deeper navigation. For the complex tasks, the academic context and task complexity increased students' standards of coherence and source processing. The academic context also fostered more navigation during simple tasks and improved search outcomes for students who struggled to understand the core topic in complex tasks. Theoretical and practical implications are discussed.

Keywords: information search, task model, task context, search behavior

Introduction

Searching for information requires individuals to construct a coherent representation of their purpose and to engage in goal-adapted strategies (Brand-Gruwel et al., 2017). Interacting with a search engine to search for information shares similarities with online reading comprehension, task-oriented reading, and purposeful reading. These activities require individuals to integrate information from multiple online documents to achieve a specific goal (Britt et al., 2018; Kannianen et al., 2019). Students can use varying online documents to solve a personal information need or a requested one. For instance, they may search the internet to determine whether to purchase a game or to find scientific articles to write an academic essay. To succeed, students must interpret contextual cues and adjust their behavior to meet the expectations of the task. Despite their frequent use of search engines, university students often struggle with critical aspects of online searching. They tend to neglect the evaluation of source reliability and credibility, even when it is crucial. Instead, they typically focus on assessing whether top-ranked search results appear relevant to their topic (Brand-Gruwel et al., 2017).

The influence of context on purposeful reading has recently received some attention (Mason et al., 2022; Rouet et al., 2017; Schoor et al., 2021, 2023). Context is a broad umbrella concept that includes features such as the environment or conditions in which individuals perform a task. It can refer to the level of authoritativeness of the task requester (Mason et al., 2022), the type of tools used (Kammerer et al., 2018) or the type of learning goal (Latini et al.,

2019). For instance, an academic context can refer to a task requested by a teacher with the objective of providing a short essay to a teacher. In contrast, a non-academic or a leisure context can refer to a task requested by a friend with the objective of answering a friend's question (Schoor et al., 2023).

As most teachers have experienced, students can greatly differ in their understanding of task instructions (Schoor et al., 2021, 2023). Empirical evidence has shown that adapting their behavior to task demands based on the context and instructions can be difficult for students (Cerdán et al., 2019; Mason et al., 2022). Analyzing task demands is a key pre-requisite for effective information search and task-oriented reading (Macedo-Rouet et al., 2024). Recent works suggest that research should further explore the criteria students apply when performing search tasks, especially since the Internet is often used as a tool for learning new information (Kammerer et al., 2018). This is critical to help students develop better search practices on the Internet, particularly when encountering unreliable information. We argue that understanding how task context features affect students' representation of task demands is crucial to improving our understanding of the criteria they use to search for information online. Few studies have examined how students' task representations are influenced by the task context and how these representations shape search behavior and outcomes. This study explored (i) how framing search tasks in terms of varying complexity within academic and non-academic contexts influences students' task model, and (ii) how these task models affect search behavior and search outcomes.

Searching for information: processes at stake and impact of task complexity

Information Searching (IS) with a search engine is a problem-solving activity in which individuals use a search engine to solve an information need (Brand-Gruwel et al., 2005; Sharit et al., 2008). Several theoretical models have been developed to define IS, and each focuses on different aspects (Von Hoyer et al., 2022). Insisting on the cognitive aspects of IS with a search engine, Sharit and collaborators (2008) emphasize that IS involves several iterative steps: (i) goal planning and initial query production, (ii) search engine result pages processing, and (iii) websites navigation. At each stage, individuals can go backwards and update their representation of their information need, reformulate their query, or select another website if they are not satisfied. Hence, IS requires complex cognitive processes such as goal planning, information evaluation and the processing of multiple documents (Kanniainen et al., 2019; Sanchiz et al., 2020). Consequently, IS outcome can greatly vary depending on individuals' strategies. Several studies demonstrate that laypeople, when using a mock-up or a real search engine to find information about complex topics, often rely on top-link heuristics to select websites (Brand-Gruwel et al., 2017; Sanchiz, Chevalier et al., 2017). Using a mock-up search engine page displaying 12 search results varying in relevance and reliability, Rodicio (2015) instructed undergraduate students to process the links for 30 minutes to prepare for a subsequent test. The results revealed that while individuals accessed more unreliable pages, they spent more time processing relevant and reliable ones, which was correlated with better test performance.

Search processes and outcomes are particularly affected by task complexity, which refers to the objective demands of a task (Campbell, 1988). Search task complexity varies as a function of the level of precision of the search goal, the number of possible paths to reach the target information, or the breadth and depth of the problem-space of the search (Byström &

Järvelin, 1995). For instance, Dosso et al. (2023) used ‘lookup tasks’ that consisted in finding simple facts vs. ‘exploratory tasks’ requiring to explore the environment to acquire new knowledge with a more complex search goal to investigate the effect of complexity on strategies and search outcome. In another study, Sanchiz, Chin and colleagues (2017) asked undergraduate students to use the *Google* search engine to solve simple search tasks consisting in fact-finding (i.e. the target information was clear, and keywords were directly provided in the task instructions), intermediate tasks that required more inferences, and complex tasks for which the instructions were less clear (i.e. ambiguous keywords were provided) so that students had to draw connections. Notably, the most complex search tasks consisted in open-ended tasks that required multiple search paths and more extensive information processing. Overall, empirical studies showed that simple search tasks, that provide clear clues to find the target answer, are often rapidly solved and result in better search outcomes. In contrast, complex tasks, that require more integration and provide fewer clues are related to deeper navigation, longer dwell time processing information and poorer search outcomes (Dosso et al., 2023; Sanchiz, Chin et al., 2017;). Similar findings are observed in task-oriented or purposeful reading tasks with multiple online documents. Using data from the PIAAC field test, Goldhammer and colleagues (2014) examined participants’ online strategies to study the relationship between time on task and performance as a function of the task demands. They demonstrated that longer time on task was negatively correlated with search outcomes for the simple reading tasks. In contrast, this correlation was either weaker or positive for more difficult reading tasks requiring complex problem-solving processes.

Altogether these findings show that the evaluation phase is critical in IS, notably for complex search tasks that require more extensive selection and processing of online documents. Students can sometimes use relevant and adapted search strategies, although they often rely on more superficial ones. Interestingly, in two studies examining how students process a search engine results page (SERP) structured like a regular SERP (study 1) or as a grid without clear top-ranked links (study 2), Kammerer and Gerjets (2017) showed that students could assess information reliability and source trustworthiness when searching for information on a controversial issue. Students could also apply different strategies depending on the organization of the SERP. In the grid condition, students were less likely to select the top-ranked links. These findings highlight that individuals use multiple criteria to evaluate and select search results. In a critical review, Vakkari (2003) pointed out that empirical studies have identified a wide range of criteria used by individuals to evaluate and select online documents. These criteria include topicality, utility, and the level of authority. Most studies he reviewed also emphasized that the selection of these criteria depends on individuals’ prior knowledge and experience. Interestingly, Vakkari (2003) concluded that information search processes and outcomes result from the interplay between the features of search tasks and individuals’ characteristics. He briefly added that “*the information requirements of the task as perceived by the user [...] is the central feature of the task influencing information searching*” (Vakkari, 2003, p 453). Although his work does not provide further insights into the impact of individuals’ perceptions of task requirements, we argue that this claim aligns with recent works showing that perception of task demands, specifically in terms of expected outcomes and behaviors, plays a crucial role in IS processing and outcomes.

Relationships between the task context, students’ task model, search behavior and search outcome.

The issue of task demands representation has been studied in research fields addressing task-oriented or purposeful reading (Britt et al., 2018; Hautala et al., 2022; Rouet et al., 2017) and online reading comprehension (Kanniainen et al., 2019). Task oriented-reading, or purposeful reading, like information searching, is a goal-directed type of reading where reading serves to fulfill a specific goal (McCrudden & Schraw, 2007). Task-oriented reading involves a complex set of inquiry processes, such as evaluating and selecting documents, as well as comprehension processes, including inference-making, coherence-building, and integration (Kanniainen et al., 2019; Macedo-Rouet et al., 2024; Schoor et al., 2023).

The cognitive representations and processes involved in purposeful reading are detailed in the Reading as a problem Solving (RESOLV) model (Rouet et al., 2017). RESOLV posits that reading always occurs in a social and a physical context that determines readers' purpose or need for reading. The task context includes several components, such as the task requester, the type of task, the environment in which the persons are performing the task, the tools they are using, the learning goals provided and the task audience (Latini et al., 2019; Mason et al., 2022; Rouet et al., 2017). Based on their understanding of the context, individuals develop a context model that shapes their interpretation of the task instructions and demands, which is referred to as the task model. This task model is developed at the beginning of the activity thanks to students' context model, prior topic knowledge, experience schemata and metacognitive knowledge, and it can be updated along the task. Hence, students' initial task model may differ, leading them to develop different goals and strategies to solve the task at hand (Britt et al., 2018; Hautala et al., 2022; Rouet et al., 2017). A prior study examined the relationships between university students' perception of the task demands and multiple documents reading by assessing the task model after task completion. Results showed that students can perceive some of the task demands related to information corroboration and integration. However, students' task model mostly included superficial processes (Schoor et al., 2021).

The effects of the task context on individuals' task model are still unclear. However, promising evidence emphasized that variations in the task context can influence individuals' expectations in terms of standards of coherence, which refers to individuals' understanding of what an acceptable level of comprehension is with regards to their goal (Van den Broeck et al., 2001). Two studies that framed a single document reading task in an entertainment context or a study context showed that readers applied more standards, made more coherence-elaborating inferences, and recalled more text information in the study context as compared to the entertainment one (Bohn-Gettler & Kendeou, 2014; Van den Broek et al., 2001). In purposeful reading tasks with multiple online documents, studies with university students showed that an academic context increases time on task, the number of switches between documents, and the selection of more expert sources, compared to a personal context (Latini et al., 2019; Schoor et al., 2023). For instance, Schoor and collaborators (2023) asked 106 university students to read multiple texts in preparation for writing a short essay. The task was framed in two contexts: (i) a university context, where participants were told that the task was assigned by a teacher and the essay would be evaluated by the teacher, or (ii) a personal context, where the task was requested by a friend and only that friend would read the essay. Findings highlighted that variations in the task context can influence students' understanding of the task demands. Consequently, students developed different expectations and search strategies. An academic context will deem more appropriate the applications of higher standards of coherence and academic norms than a personal context, which should result in more source evaluation and integration processes for instance (Schoor et al., 2023). Overall, the few studies examining the

interplay between the task context and the task demands seem to point out that the task context and the task demands can influence students' task model and affect the standards of coherence applied (Rouet et al., 2017; Schoor et al., 2023).

Bridging the theoretical frameworks of information search and task-oriented/purposeful reading, the present state-of-the-art pinpoints that to develop adapted standards and engage in effective search strategies, students must first ensure they understand what the task requires in terms of behaviors and outcomes. In other words, students should pay particular attention to the features of the search tasks they are requested to solve to form a relevant task model. Indeed, studies have shown that the task model can be related to online reading processes for undergraduate students (Cerdán et al., 2013; Schoor et al., 2021). However, one study failed to find significant relationships between aspects of the task model related to multiple documents processing and actual behavior (Schoor et al., 2023). Despite this, no prior studies have investigated the impact of the quality of students' task models on online information processing in IS or task-oriented reading.

Examining students' task model quality highlights the need for an appropriate assessment tool. The representation of task demands is often explored through the lens of students' goals and sub-goals. In a study about self-regulated learning with multiple documents, McCardle and collaborators (2017) introduced the TASC framework to examine students' goal quality. They argue that a high-quality goal should mention four components: Time, Actions, Standards and Content, that cover the complex layering of processes at stake when using multiple online documents to learn. Time refers to students' will to set short-term goals. Actions reflect students' cognitive processing. Standards refers to students' will to judge their own progress and reflect upon their behavior while completing the task. Content refers to the topic and issue at stake in the task. In their study, students had to set a studying goal for the upcoming week, and to reflect on their goal attainment from the previous week. They measured quality goals by assessing each of the four TASC elements. They showed that students rarely included specific information regarding actions, standards, or content.

The current study

This study examined (i) the extent to which variations in the search task context influenced students' representation of the task demands (*i.e.* students' initial task model) for simple and for complex search tasks, and (ii) how students' initial task model could influence search behavior and outcomes when using a search engine. To this aim, university students completed two simple and two complex search tasks, each framed within an academic or non-academic context. To measure students' initial task model, participants read each task instruction and, before searching the Internet, answered two questions designed to describe what the task required them to do.

Our first research question (RQ1) investigated whether students' task models change as a function of the task context and complexity. Relying on the TASC framework (McCardle et al., 2017), we hypothesize that elaborate task models include components related to complex cognitive processes, such as planning (*e.g.* the Time component) and information evaluation (*e.g.* the Standards component). We expected that the academic context would activate higher standards of coherence and academic norms, resulting in a more coherent task model that includes planned and self-regulated strategies (Rouet et al., 2017; Schoor et al., 2021, 2023, H1a). Additionally, we assumed that the students would develop a more elaborate task model for complex tasks compared to simple tasks (H1b).

Our second research question (RQ2) examined the impact of task context and students' task model on search outcomes and search behavior. In line with prior works showing that simple and complex tasks greatly varied in terms of task demands (Dosso et al., 2023; Sanchiz, Chin et al., 2017), we addressed RQ2 by considering simple and complex tasks independently. Indeed, the present study used simple and complex tasks like Sanchiz, Chin and colleagues (2017) that are comparable with the simple functional reading tasks and the more complex problem-solving tasks designed by Goldhammer and colleagues (2014). These studies all showed that such simple and complex tasks require different processes, which may not be directly comparable. A summary of the impact of task complexity on our dependent variables can be found in Appendix 1. Since the focus of our study was on understanding how students perceive task demands, we argue that combining simple and complex tasks would have introduced unnecessary ambiguity. In addition, examining the specific impact of the search tasks complexity on search behavior and outcomes outside the scope of this study, making it unnecessary to consider simple and complex tasks together. Regarding the impact of the task model on search behavior and outcomes, we expected that a more elaborate task model would improve search outcomes, behavior, and source recall (H2a). Conversely, a more superficial initial task model would likely reduce the time spent processing information across different sources, instead leading to more time spent on search engine result pages (H2b). Finally, we hypothesized that the academic-related context would increase the time spent processing information, the number of webpages consulted and search outcomes, compared to non-academic related tasks (H3).

Method

Sample

Ninety first-year Psychology university students enrolled in this experiment, (age range 17-27 years, $M = 18.51$ $SD = 0.84$, 94.81% female). They were recruited via an online application from the Psychology department and received a credit course for participation. The sample size was estimated based on prior studies examining the impact of task model on reading comprehension (Cerdán et al., 2019; Schoor et al., 2023) and recruitment feasibility. The study followed the ethical guidelines of the authors' Faculty and the French General Rule about personal data protection.

Design

There were two independent within-subject variables. Context was operationalized as the level of academic-relatedness of the search tasks topics (academic-related or non-academic related). Complexity consisted in simple fact-finding and complex tasks.

Material

Search tasks

The academic-related and the non-academic related tasks were designed by applying the two following criteria identified in prior studies: academic tasks had to tackle a topic relevant

to the participants' university curriculum and they had to be requested by an authoritative source. As participants were Psychology students, Educational Psychology was used as a proxy for the academic-related task context because this topic was part of their curriculum. Manga, which is a Japanese comic book genre, was used as a proxy for the non-academic related task context because it was not related to the participants' university curriculum.

Simple and complex tasks were designed based upon the following criteria. Simple tasks were tasks that could be quickly solved, that provided clear keywords about the target information and how to access it. Simple tasks were fact-finding tasks that provided all information necessary to find the answer. Participants had to select the relevant keywords in the task instructions, use them in a query and process the SERPs to find the answer. Complex tasks had to be less clear, they had to require more problem-solving processes and web navigation, and the type of outcome needed to be more demanding. Hence, the two complex tasks required (i) to process multiple webpages to understand opposite views on a topic (autism or the diversity of Mangas), (ii) to collect and integrate information about these conflicting topics, and (iii) to provide a short written argumentative synthesis of the two viewpoints. The task demands of the simple and the complex tasks were held constant across the two task contexts. See Table 1 below for the task instructions.

Table 1. *Task instructions in English (instructions originally in French)*

Tasks	Non-academic related topic	Academic related topic
Simple	When did Jirô Taniguchi receive the Knight of the Order of Arts medal?	As part of your studies, you are interested in pedagogical practice. Maria Montessori is a great figure of pedagogical innovation in France. When did Maria Montessori receive France's <i>Legion of Honor</i> medal?
Complex	Charlie, 22, likes Mangas. He would like to introduce Mangas to his relatives, but they associate these types of books with violence, immature stories, and stereotypical characters. In a short synthesis, search the internet to define whether Mangas can be a diversified literary genre.	As part of your studies, you are asked to search for information to produce a synthesis about the origin of autism. Scientific publications are contradictory, some argue that autism is caused by genetic factors while other point that environmental factors. Search the internet to produce a brief synthesis about the causes of autism which could be used as a basis for an oral presentation.

Data: search behavior, search outcome and source recall

Two search behavior indicators were used: (i) the total number of webpages consulted, and (ii) the proportion of time spent on SERPs (as opposed to the proportion of time spent processing webpages).

Search outcome was assessed with answer accuracy. For simple fact-findings tasks, one point was granted for the expected target information and zero for all other answers. Complex tasks required participants to: (i) gather and process information, (ii) describe in a synthesis the two viewpoints to present a conclusion about the conflicting topic. Zero points were granted for a yes/no answer that included no argument. One point was granted for partial answers that included detailed arguments about one of the two viewpoints. Two points were granted for complete answers that detailed information about the two viewpoints and/or some reasoning (such as a critical comment on the source used, or an attempt to explain the conflict). Each answer was coded independently. Ten answers were used as training cases by two raters, which resulted in good reliability ($\kappa > .82$). The remaining answers were coded by one rater.

Source recall was assessed with one open-ended question that asked students to write down as many details as they could recall about the information source they used to find the target answer (e.g. website name, website topic).

Initial task model measurement and coding scheme

Participants answered two open-ended questions before completing the task: “Explain in your own words what this search task requires you to do” and “In your opinion, what does a person should do to find the target information?”. Authors created a coding system based on theoretical considerations and prior works to discriminate elements related to search processes or to the search topic (also called the core). Students’ initial task model was divided into four sub-components: « Timeframe », « Actions », « Standards », « Core » (McCardle et al., 2017). For each component, a score was attributed. See Table 2 for examples. One category corresponded to the core (*i.e.* topic) and three categories were related to processes:

- Timeframe: referred to temporal planification or regulation strategies
- Actions: three general types of actions involved in information search with a search engine were considered: querying, navigating and information processing.
- Standards: referred to information evaluation, source processing, corroboration need, comparison of websites content, relevance assessment with regards to the search goal.
- Core: referred to the number of main ideas/keywords involved in the task topic (up to three for the simple tasks and two for the complex ones). Results were then transformed into percentage to ensure comparability between simple and complex tasks.

Table 2. *Examples of answers and coding criteria for participants’ task model (translated from French to English)*

Examples	Timeframe	Actions	Standards	Core
Search with keywords connected to the topic, then navigate through multiple relevant and reliable websites.	1	2	1	0
In a search engine, I will enter " Jirô Taniguchi" and " knight of Arts and Letters". In general, I click on the first link that appears if it seems relevant. Otherwise, I choose the next one and so forth.	1	3	1	2
I will enter my query and search on the first websites	0	2	0	0
I will search "year Mr. Jirô Taniguchi Knight of Arts and Letters" on a search engine.	0	1	0	3

Each student’s answer was coded as follows. Timeframe and standards were coded 1 if participants’ answers mentioned elements that referred to these criteria. Other answers were coded 0. Actions were coded on a 3-point scale: one point was granted per type of possible action mentioned (querying, navigating and information processing). Finally, for the core, we counted the number of main relevant concepts that were included in participants’ answers (three for simple tasks and two for complex tasks). Mean score for each component were calculated based on the two questions used to assess the task model. Scores were then transformed into percentage for the analysis to ensure comparison between simple and complex search tasks.

Two raters coded 15 answers together as training cases. Data from two participants were dropped as they did not provide any answer. Disagreements were solved and modifications of the coding system were made. Next, all answers for one simple and one complex search task were coded independently by the two raters as a second wave of training cases. Table 3a and 3b below summarize inter-rater reliability.

Table 3a. Summary of inter-reliability testing for the training cases for the non-academic related simple search task.

Coding category	Percentage of coding agreements	Cohen's κ
Time frame	95	.84
Actions	80	.53
Core	88	.83
Standards	98	.91

Table 3b. Summary of inter-reliability testing for the training cases for the academic related complex search task.

Coding category	Percentage of coding agreements	Cohen's κ
Time frame	88	.64
Actions	69	.38
Core	96	.94
Standards	93	.86

Overall, inter-reliability was very good for standards (Cohen's $k > .86$) and for core (Cohen's $k > .83$). Timeframe inter-reliability was average for the complex academic related search task ($k = .64$). The inter-reliability coding of the Actions component was also low ($k = .53$ for the simple non-academic related task and $k = .38$ for the academic related task). All disagreements were identified and solved, resulting in a second wave of modifications to clarify ambiguities. As the two raters reached consistent agreement (both $k > .79$), they finally coded half of all answers independently.

Procedure

The experiment was run in groups of ten, in a classroom setting with laptops and the *Google* search engine. Each participant worked on a 15-inch screen laptop and used the *Chrome* browser and *Google* search engine to solve the search tasks. Logfiles were extracted via the *Chrome History View* app. The experiment was run within a week, so that no major changes that could have influenced the results retrieved in the search engine results pages appeared. Participants first completed an informed written consent. Then, participants provided socio-demographic information and internet habits with an online questionnaire (Limesurvey).

Next, participants completed the four search tasks in pseudo-randomized order with the following procedure. (1) Participants first read the task instructions, (2) they answered two open-ended questions assessing their task model, and (3) they rated on a 5-point Likert scale whether they knew the answer. No participants stated that they knew the answer to any of the four tasks. Next, (4) participants searched for the answer on the Internet. Once participants were done searching, they called the experimenter who closed the navigator and switched back to the questionnaire. (5) Participants entered their answer and (6) they completed questions assessing satisfaction, and certainty of answer. Finally, (7), participants had to recall as much information

as they could about their information source. Students were told that they had 60 minutes maximum to complete the four tasks.

Data analysis

Results were analyzed with Jamovi (version 2.2.5), a free software built on R. To answer RQ1, we conducted mixed ANOVAs with task complexity and academic relatedness of the search topic (*i.e.* task context) as independent variables and the four components of the task model as dependent variables. To answer RQ2, linear mixed-effect models were used with the Gamlj Package and bobyqa optimizer to examine whether the task model and the task context could predict search outcome and search strategies in simple and complex tasks. As it was not the focus of the present study, we do not present results regarding the impact of task context and task complexity on search behavior and search outcomes. Yet, readers may find a summary of findings in Appendix 1.

Results

RQ1: Does students' initial task model change as a function of task context and task complexity?

First, a descriptive screening of answers showed that around 84% of participants did not mention timeframes elements in their initial task model. Around 95% of participants mentioned at least one type of action. More specifically, 69.76% mentioned one type of action, 26.04 % mentioned two different types of actions and 4.2% mentioned all three types of possible actions (*i.e.* querying, navigation and information processing). Likewise, at least one core element was included in 63.35% of answers. Standards were also rarely mentioned in participants' initial task model: 69.71 % of participants mentioned 0 standards.

Task model timeframe

Academic relatedness had no impact on timeframe elements, $F(1,84) = .01, p = .91, \eta^2_p < .001$. 83.53% of participants and 81.49% in the non-academic related topic did not mention timeframe elements. Complexity was not significant $F(1,84) = .13, p = .72, \eta^2_p = .002$. The interaction reached significance, $F(1,84) = 20.62, p < .001, \eta^2_p = .19$. For simple tasks, more timeframe elements were mentioned for the non-academic related topic than for the academic one ($t(84) = 3.34, p = .008$). For complex tasks, fewer timeframe elements were mentioned for the non-academic related topic than for the academic one ($t(84) = -2.69, p = .052$). See Table 4 below.

Table 4. *Frequency of the number timeframe elements reported in students' initial task model*

Number of timeframe elements	Non-academic related topic		Academic related topic	
	Simple tasks	Complex tasks	Simple tasks	Complex tasks
0	75,29%	91,76%	91,86%	76,74%
0,5	21,17%	6,98%	8,13%	17,44%
1	3,52%	1,26%	0%	5,81%

Task model actions

Academic relatedness and task complexity had no impact on action elements, $F(1,84) = 1.28, p = .26, \eta^2_p = .015$ and $F(1,84) = .73, p = .40, \eta^2_p = .009$. The interaction was significant, $F(1,84) = 8.54, p < .004, \eta^2_p = .09$. For simple tasks, context had no impact ($t(84) = 1.03, p = 1.00$), whereas for complex ones, action elements were less mentioned for the non-academic related topic than for the academic one. ($t(84) = -3.20, p = .01$). See Table 5 below.

Table 5. *Frequency of the number action elements reported in students' initial task model.*

Number of action elements	Non-academic related topic		Academic related topic	
	Simple tasks	Complex tasks	Simple tasks	Complex tasks
0	4,70%	2,35%	1,17%	2,35%
0,5	3,52%	3,52%	9,41%	2,35%
1	61,17%	75,29%	65,88%	56,47%
1,5	16,47%	9,41%	14,11%	21,17%
2	11,76%	9,41%	8,23%	15,29%
2,5	1,176%	0	1,176%	2,35%
3	1,176%	0	0	0

Task model standards

Fewer standards were mentioned for the non-academic related topic as compared to the academic one, $F(1,84) = 32.5, p < .001, \eta^2_p = .28$. Respectively 80% and 58.25% of participants mentioned standards in the academic related and in the non-academic related topics. Fewer standards were mentioned for simple tasks as compared to complex ones, $F(1,84) = 20.5, p < .001, \eta^2_p = .20$. Finally, the interaction was significant, $F(1,84) = 39.3, p < .001, \eta^2_p = .32$. For simple tasks, academic relatedness of the topic had no impact ($p = 1.00$). In contrast, for complex tasks, fewer standards were mentioned for the non-academic related topic than for the academic related one. ($t(84) = -7.78, p < .001$). See Table 6.

Table 6. *Frequency of the number standard elements reported in students' initial task model.*

Number of standards elements	Non-academic related topic		Academic related topic	
	Simple tasks	Complex tasks	Simple tasks	Complex tasks
0	78,82%	81,17%	78,23%	61,17%
0,5	14,11%	16,47%	15,29%	20%
1	7,05%	2,35%	6,47%	18,82%

Task model core

As a reminder, the task model core assessed to what extent the participants provided details about the core topic of each search task (*i.e.* the level of details of the topicality). We evaluated whether participants provided extensive details about the core topic of the task (*i.e.* all main keywords or synonyms), some details (*i.e.* half of all possible main keywords or synonyms) or no details at all (*i.e.* no main keywords or synonyms). Academic relatedness had no impact on core elements, $F(1,84) = 2.54, p = .11, \eta^2_p = .029$. Task complexity reached significance $F(1,84) = 5.91, p = .017, \eta^2_p = .066$. For the simple tasks, participants provided less details about the core of the search task (about 44% of the core topic was mentioned) than in the complex search tasks (about 52% of the core topic was mentioned) (respectively $M\% = 44.4 SE = 2.62$ and $M\% = 51.9 SE = 3.36, t(84) = -2.43, p = .017$). Finally, the interaction was not significant, $F(1,84) = 2.90, p = .09, \eta^2_p = .033$. See Table 7 below.

Table 7. Frequency of the number core elements reported in students' initial task model.

Number of core elements	Non-academic related context		Academic related context	
	Simple tasks	Complex tasks	Simple tasks	Complex tasks
0	22,35%	16,47%	20%	29,41%
0,5	5,88%	11,76%	7,05%	8,23%
1	12,94%	25,88%	15,29%	30,58%
1,5	23,52%	20%	25,88%	8,23%
2	20%	25,88%	15,29%	23,52%
2,5	7,05%	0%	9,41%	0%
3	8,23%	0%	7,05%	0%

Summary of results for RQ 1

Partially as expected, academic relatedness increased initial task model standards, but it had no impact on timeframe and actions. Contrary to hypothesis 1a, task model core was more detailed in the simple as compared to the complex tasks. Finally, complexity increased standards and core elements in students' initial task model, but it had no influence on timeframe and actions. The interaction between task context and complexity showed that for complex tasks, the academic relatedness of the search topic increased timeframe, actions, and standards in students' task model, which tends to validate H1b.

RQ 2: relationships between task context, students' initial task model, search behavior and search outcome for simple tasks

Preliminary analyses for the present study also confirmed that the simple and complex tasks varied in terms of task demands (see Appendix 1). Hence, we decided to address RQ 2 by considering simple and complex tasks independently. Descriptive measures and zero-order correlations among the measures for simple tasks are presented in Table 8. There was a positive correlation between core and search outcome ($r = .27, p < .001$) and a negative correlation between standards and search outcome ($r = -.15, p = .04$). There was also a positive correlation between standards and the number of webpages consulted ($r = .22, p < .01$) and between standards and source recall ($r = .15, p < .05$). Table 9 provides Means and SEs for search behavior, search outcome and source recall for simple tasks, as a function of task context. Appendix 2 provides Means and SDs for the time spent on the search engine results pages and the time spent on webpages. As a reminder, these indicators were computed as the proportion of time spent on the search engine results pages.

Table 8. Zero-order correlations among variables for the experimental group ($N = 90$) for the simple tasks.

	1	2	3	4	5	6	7	8
1 Timeframe TM	-							
2 Actions TM	.49***	-						
3 Standards TM	.36***	.59***	-	-.14				
4 Core TM	.04	.12	-.14	-				
5 Number of webpages consulted	.04	.07	.22**	.02	-			

6 Proportion of time on SERPs	-.12	-.13	-.04	.04	-.48***	-		
7 Source recall	-.04	.11	.15*	.001	.36***	-.05	-	
8 Search outcome	-.004	.04	-.16*	.27***	.10	-.01	.05	-

Note. * $p < .05$, ** $p < .01$ *** $p < .001$, TM = Task Model

Table 9. Means and (SE) for simple tasks.

	Academic relatedness of the task context	
	Non-academic related topic	Academic related topic
Proportion of time on the search engine results pages (%)	57.8 (4.15)	37.7(3.65)
Number of webpages visited	.88 (.11)	2.10 (.22)
Source recall (max =3)	1.36 (0.06)	1.66 (0.18)
Search outcome	91.8 (3.00)	87.1 (3.66)

A linear mixed model was fitted (estimated using REML and bobyqa optimizer) (Powell, 2009) to predict search outcome and search behavior with: the four components of the task model, task context (non-academic related topic as compared to the academic related one), and their respective interaction as predictors. The model included Subject as random effects. Only significant findings are reported below.

Search outcome

We expected that a more elaborate initial task model would increase search outcome (H2). However, no significant effects were observed (all $p > .05$). Likewise, when considering the students' task model and the task context, no effect of the task context was observed on search outcomes (H3).

Search behavior: number of webpages consulted, and proportion of time spent on the search engine results pages

Regarding the number of webpages consulted, the model's total explanatory power was conditional $R^2 = 0.15$, and the part related to the fixed effects alone (marginal R^2) was 0.31. The model's intercept was .34 (95% CI [0.001, 1.773]). Students who had higher initial task model standards consulted more webpages in the academic-related topic than in the non-academic related one ($B = 1.43$, $t = 3.03$, $p = .003$, 95% CI [.506, 2.36]). Task context also reached significance ($B = .87$, $t = 4.34$, $p < .001$, 95% CI [.476, 1.259]): the academic related topic increased the number of webpages consulted. See Table 10.

Table 10. Fixed effects parameters estimates for simple tasks number of webpages consulted

Name	Estimate	SE	95% CI	<i>t</i>	<i>p</i>
Intercept	1.31	0.12	1.08, 1.54	11.12	< .001
Task context	0.87	0.20	0.48, 1.26	4.34	< .001
Timeframes	1.10	0.69	-0.25, 2.44	-1.36	0.11
Actions	-0.42	0.31	-1.03, 0.19	-1.36	0.18
Core	0.05	0.12	-0.19, 0.29	0.40	0.69
Standards	1.43	0.47	0.51, 2.36	3.03	0.003
Task context *Timeframes	1.80	1.33	-0.80, 4.41	1.36	0.18
Task context *Actions	-0.48	0.61	-1.67, 0.72	-0.78	0.43
Task context *Core	-0.22	0.24	-0.68, 0.24	-0.92	0.36

Task context *Standards	0.96	0.94	-0.87, 2.80	1.03	0.30
-------------------------	------	------	-------------	------	------

Note: reference level for the task context was the non-related academic domain.

For the proportion of time spent on the search engine results pages, the model's total explanatory power was conditional $R^2 = 0.11$, and the part related to the fixed effects alone (marginal R^2) was 0.268. The model's intercept was 279 (95% CI [0.001, 680]). Within this model, only the task context reached significance ($B = -19.70$, $t = -3.27$, $p = .002$, 95% CI [-31.50, -7.91]). In the academic related topic, participants spent a lower proportion of time on the search engine results pages than in the non-academic related topic. See Table 11.

Table 11. Fixed effects parameters estimates for simple tasks proportion of time spent on SERPs

Name	Estimate	SE	95% CI	t	p
Intercept	50.59	3.52	43.69, 57.50	14.36	< .001
Task context	-19.70	6.02	-31.50, -7.91	-3.27	0.002
Timeframes	-12.88	20.42	-52.91, 27.14	-0.63	0.53
Actions	-9.55	9.53	-28.22, 9.13	-1.00	0.32
Core	2.03	3.70	-5.23, 9.29	0.55	0.58
Standards	14.09	14.10	-13.55, 41.73	0.99	0.32
Task context *Timeframes	52.77	39.77	-25.17, 130.71	1.33	0.19
Task context *Actions	11.56	18.96	-25.61, 48.72	0.61	0.54
Task context *Core	-3.17	7.13	-17.14, 10.80	-0.45	0.66
Task context *Standards	10.04	27.91	-44.66, 64.75	0.36	0.72

Note: reference level for the task context was the non-related academic domain.

Source recall

Regarding source recall, the model's total explanatory power was conditional $R^2 = 0.29$, and the part related to the fixed effects alone (marginal R^2) was 0.293. The model's intercept was 2.21 (95% CI [0.001, .338]). Within this model, only task context reached significance ($B = 1.82$, $t = 7.64$, $p < .001$, 95% CI [1.36, 2.29]). See Table 12.

Table 12. Fixed effects parameters estimates for simple tasks source recall

Name	Estimate	SE	95% CI	t	p
Intercept	2.26	0.12	2.03, 2.49	18.93	< .001
Task context	1.82	0.24	1.35, 2.29	7.63	< .001
Timeframes	0.97	0.75	-0.50, 2.44	1.29	0.20
Actions	0.32	0.34	-0.35, 0.99	0.92	0.36
Core	-0.06	0.13	-0.32, 0.20	-0.47	0.64
Standards	0.59	0.52	-0.43, 1.62	1.13	0.26
Task context *Timeframes	2.28	1.50	-0.66, 5.22	1.52	0.13
Task context *Actions	-0.14	0.68	-1.48, 1.20	-0.20	0.84
Task context *Core	-0.07	0.26	-0.59, 0.45	-0.27	0.78
Task context *Standards	1.59	1.05	-0.46, 3.64	1.52	0.13

Note: reference level for the task context was the non-related academic domain.

Summary of results for RQ 2 with simple tasks

Unexpectedly, no impact of students' initial task model and task context was observed on simple search tasks outcomes (H2 and H3). However, higher standards in the students' initial task model increased the number of webpages visited. The academic related context increased the number of webpages visited and source recall, but it reduced the proportion of time spent on the search engine results pages.

RQ 2: relationships between students' initial task model, search behavior and search outcome for complex tasks

Descriptive measures and zero-order correlations among the measures for the complex tasks are presented in Table 13. Actions were correlated to the proportion of time spent on SERPs ($r = .24, p = .03$), and standards were also related to the number of webpages visited ($r = .23, p = .03$). Core components were not correlated to any variables. In addition, search outcome was related to the proportion of time spent on SERPs ($r = -.23, p = .02$) and source recall ($r = .23, p = .015$). Source recall was positively related to timeframes ($r = .24, p = .02$), actions ($r = .48, p < .001$), and standards ($r = .60, p < .001$). Table 14 provides Means and SEs for search behavior, search outcome and source recall for complex tasks, as a function of task context.

Table 13. Zero-order correlations among variables for the experimental group ($N = 90$) for the complex tasks

	1	2	3	4	5	6	7	8
1 Timeframe TM	-							
2 Actions TM	.52***	-						
3 Standards TM	.19	.64***	-					
4 Core TM	.17	.12	-.20	-				
5 Number of webpages consulted	-.06	.02	.23*	-.01	-			
6 Proportion of time on SERPs	.16	.24*	.02	.15	-.22*	-		
7 Source recall	.24*	.48***	.60***	-.11	.02	.08	-	
8 Search outcome	.17	.28*	.20	.08	.17	-.23*	.23*	-

Table 14. Means and (SE) for complex tasks.

	Academic relatedness of the topic	
	Non-academic related topic	Academic related topic
Proportion of time on the search engine result pages (%)	32.0 (3.28)	23.4 (2.97)
Number of webpages visited	2.20 (.21)	2.54 (.37)
Source recall (max =3)	0.70 (0.07)	1.87 (.10)
Search outcome (%)	52.4 (4.34)	60.0 (4.08)

A linear mixed model was fitted (estimated using REML and bobyqa optimizer) (Powell, 2009) to predict search outcome and search behavior with the four components of the task model, task context and their respective interaction. The model included Subject as random effects. Only significant findings are reported below.

Search outcome

Regarding search outcome, the model's total explanatory power was substantial (conditional $R^2 = 0.32$), and the part related to the fixed effects alone (marginal R^2) was 0.08. The model's intercept was 376 (95% CI [132, 796]). Within this model, the only significant fixed effects observed was the interaction between the task context and the task model core (B

= -19.70, $t = -2.46$, $p = .015$, 95% CI [-35.37, -4.02]). Students who had a low initial task model core obtained poorer search outcome in the non-academic related topic as compared to the academic related one (respectively $M_{\%} = 40.0$ $SE = 7.00$ and $M_{\%} = 64.0$ $SE = 6.29$). See Table 15 below for fixed effects parameters estimates.

Table 15. Fixed effects parameters estimates for complex tasks search outcome

Name	Estimate	SE	95% CI	t	p
Intercept	53.64	3.80	46.19, 61.08	14.11	< .001
Task context	9.50	6.38	-3.00, 22.00	1.48	0.14
Timeframes	11.58	14.37	-16.58, 39.74	0.80	0.42
Actions	12.53	9.57	-6.23, 31.28	1.31	0.19
Core	2.24	4.52	-6.62, 11.09	0.49	0.62
Standards	-12.33	12.35	-36.53, 11.87	-0.99	0.32
Task context *Timeframes	17.83	28.35	-37.74, 73.40	0.63	0.53
Task context *Actions	-.13	18.46	-36.32, 36.05	-0.007	0.99
Task context *Core	-19.70	8.00	-35.37, -4.02	-2.46	0.02
Task context *Standards	12.82	24.39	-34.98, 60.61	0.52	0.60

Note: reference level for the task context was the non-related academic domain.

Search behavior: number of webpages consulted, and the proportion of time spent on the search engine results pages

Regarding search behavior, no significant results were found for the total number of webpages visited and the proportion of time spent on the search engine results pages. See Table 14 for Means and SE.

Source recall

Regarding source recall, the model's total explanatory power was conditional $R^2 = 0.38$, and the part related to the fixed effects alone (marginal R^2) was 0.50. Within this model, the positive impact of standards was significant ($B = 1.029$, $t = 2.66$, $p = .009$, 95% CI [.270, 1.787]) as well as the interaction between task model standards and task context ($B = 2.08$, $t = 2.71$, $p = .007$, 95% CI [.578, 3.586]). Students who had higher initial task model standards better recalled source information in the academic-related topic than in the non-academic related topic (respectively $M = .70$ $SE = .39$ and $M = 1.75$ $SE = .17$). See Table 16 for Fixed effects parameters estimates.

Table 16. Fixed effects parameters estimates for complex tasks source recall.

Name	Estimate	SE	95% CI	t	p
Intercept	0.821	0.115	0.595, 1.05	7.11	< .001
Task context	0.239	0.204	-0.161, .64	1.17	0.24
Timeframes	0.406	0.451	-0.477, 1.29	0.90	0.37
Actions	0.326	0.299	-0.260, .91	1.09	0.28
Core	-0.131	0.139	-0.404, .14	-0.94	0.35
Standards	1.029	0.387	0.270, 1.79	2.66	0.01
Task context *Timeframes	0.440	0.892	-1.309, 2.19	0.49	0.62
Task context *Actions	0.416	0.583	-0.726, 1.56	0.71	0.48
Task context *Core	-0.196	0.255	-0.696, .30	-0.77	0.44
Task context *Standards	2.082	0.767	0.578, 3.59	2.71	0.00

Note: reference level for the task context was the non-related academic domain.

Summary of results for RQ 2 for complex tasks

Unlike expected, the initial task model quality had no main impact on the search outcome for complex tasks. However, an interaction between the task model core and the task context showed that the students who elaborated a poor core in their initial task model had better search outcome for the academic related topic as compared to the non-academic related one. Regarding the number of webpages consulted and the proportion of time spent on the search engine results pages, no significant impact of the task model was observed. However, higher standards in the students' initial task model increased source recall, and this effect was qualified by an interaction. When participants had stronger standards in their initial task model, they better recalled information sources in the academic-related topic as compared to the non-academic related one. No differences were observed when participants had poor standards. These patterns tend to partially confirm H2a. Unlike expected (H3), when considering the students' task model and the task context, the task context did not significantly impact search outcome, the number of webpages visited, the proportion of time on the search engine results pages, and source recall.

Discussion

The present study had two main goals: (i) examine how task context and task complexity impact university students' initial task model in information search tasks with a search engine, and (ii) examine to what extent students' initial task model and task context could predict search outcome, search behavior and source recall for simple and complex search tasks. Participants had to solve two simple, and two complex tasks framed in an academic-related context and in a non-academic related context. The task context was operationalized as (i) a task topic related or not related to the participants' university curriculum, and (ii) the level of authoritativeness of the task requester. Hence, Psychology was used as a proxy for the academic-related context and Mangas for the non-academic related context. The academic-related tasks were framed as fictitious tasks that could be a part of the participants' studies whereas the non-academic related tasks had no such framing. Simple tasks required to select the useful keywords provided in the task instructions and to process the results pages to find the target information. The complex tasks required to understand two conflicting views on a topic and to provide a short synthesis. To the best of our knowledge, this is the first time that the relationships between task context, task complexity and students' task model are studied in information search with a search engine.

RQ1: Does students' initial task model change as a function of task context and task complexity?

Students' initial task model mostly included core and action components, but it less frequently included timeframes and standards elements. First, core and actions components are most likely the least cognitively demanding components of the task model, particularly for students familiar with information searching, and the easiest components to verbalize. Indeed, the task instructions provided useful keywords (core components) and the actions components could be easily mentioned (e.g., enter a query, see a website, pay attention to the source etc.). Hence, this pattern of results could be related to students' difficulties to verbalize higher-order automatized processes (Dominowski, 1988). However, to nuance such interpretation, prior findings showed that students tend to primarily assess the topic relevance of the search results with regards to their search need topic (Brand-Gruwel et al., 2017; Kammerer & Gerjets, 2017; Vakkari, 2003). Students tend to predominantly allocate their resources to forming a coherent

representation of the core content of the task, and a rough representation of what should be done by activating stored schemata in memory (*i.e.* surface-level actions such as producing a query and reading websites). Future research is needed, as students' representations of the utility of the search engine or their awareness of how search engines work could play a role (Kammerer & Gerjets, 2017; Kammerer et al., 2018). For instance, the more students trust the search engine, the less they may feel the need to allocate resources toward forming a coherent task model.

Secondly, findings emphasized that students may struggle to include in their initial task model the higher-order requirements of search tasks solely with the task context, the task instructions, and prior experiences or knowledge. Understanding to what extent a task will require deep-level processing such as planification and evaluation can first require to try to solve the task. Hence, theoretically, forming an elaborate initial task model includes *a minima* task expectancy, such as the perceived difficulty of the task at hand or the level of scrutiny one will need to engage (Rouet et al., 2017). This degree of scrutiny depends on prior knowledge, prior experience, motivation and on the features of the search tool (Naumann, 2019). Future works could investigate how the initial task model is updated early in the task, particularly after students have consulted documents.

Findings also confirmed that individuals may considerably differ regarding their interpretation of the task instructions as a function of the task context and complexity (McCrudden & Schraw, 2007; Rouet et al., 2017). As expected (H1a and b), the academic-related context and the complex task do tend to enhance students' expectations in terms of coherence and to lower students' attention towards content topicality (Schoor et al., 2021, 2023). Students can use clues related to the academic relatedness of the topic and the complexity of the task instructions to adapt their task model standards and actions.

RQ 2: relationships between the task context, students' initial task model, search behavior, source recall, and search outcome

For the simple tasks, students' initial task model standards were associated with more navigation, which most likely reflected their need to corroborate information and confirm their answer. This is consistent with prior studies showing that skilled comprehenders can seek out additional information and avoid hasty decision-making when the task demands require more information corroboration (Hahnel et al., 2018). No other main impact of the initial task model was found for the simple tasks, which could be related to the level of easiness of the simple search tasks. Even if students had a poor representation of the simple tasks demands, they could rely on the search engine to retrieve the correct answer. Hence, for simple search tasks students' searching abilities or basic reading comprehension skills may matter more than their initial task model (Naumann, 2019). We also cannot rule out the possibility that updated versions of students' task model are more important than their initial task model. Indeed, as discussed earlier, students may first try to retrieve the target information and then adjust their understanding of the standards and the strategies required (Naumann, 2019). However, we did not measure updated versions of the task model, so this question is left open.

Of note, for the simple tasks, results showed that the academic related context increased the proportion of time spent navigating webpages as opposed to the proportion of time spent on the results pages. These findings emphasize that an academic-related context can highlight the importance of avoiding shallow information processing of the search engine results (Hahnel et al., 2018). In addition, the academic-related context also improved search outcome for students who struggled to understand the core topic of the complex tasks.

For the complex tasks, results showed that students who had a poor initial task model core had better search outcome for the academic-related as compared to the non-academic

related tasks. The academic context may have led students to process information more deeply, which probably fostered a better understanding of the task topic and helped students cope with their low initial task model core and improved their search outcome. In addition, initial task model standards increased source recall for complex tasks. In line with prior studies (Latini et al., 2019; Schoor et al., 2021,2023), students perceived the higher demands of the complex tasks in terms of sourcing, information corroboration and integration. Thus, they paid more attention to sources to ensure they had found reliable information. This pattern was particularly pronounced in the academic related task. Students who had higher standards recalled more information source for the academic related complex tasks than for the non-academic related ones.

Overall, results call for future research regarding: (i) the impact of updated versions of the task model, and (ii) the assumption that a minimal processing of the documents can be required to elaborate a task model that supports relevant search behavior (Naumann, 2019). Findings emphasize that students can include in their initial task model the deep-level multiple documents comprehension processes required by the search task (Schoor et al., 2021). However, prior knowledge, prior experience and student's initial understanding of the task context and instructions do not seem sufficient to help them to fully grasp deep-level regulation processes. The demands related to regulation and metacognitive processes seem to be included in updated versions of the task model. Such a pattern may provide interesting insights for designing interventions or support tools aimed at helping students develop deep-level regulatory strategies in information search or task-oriented reading.

Limitations

First, sample representativity is a limitation because we needed participants with the same curriculum to ensure the (non-)relatedness of the search topics to their studies. We only recruited Psychology students, which mostly consisted of a female sample. In addition, the task context was confounded with the search topic, as different topics were used to operationalize the nonacademic vs. academic contexts. This confound may have influenced the study outcomes, as the perceived importance or relevance of a topic could lead participants to engage in deeper reflective processing, thereby contributing to a more complex representation of the task demands. For instance, a participant might find a particular topic important due to a personal passion or its perceived utility, regardless of the academic-related context. This could skew the findings, especially for complex tasks associated with an authoritative audience. Future studies should disentangle task context from topic characteristics, potentially by manipulating the level of authoritativeness of the task requester and the audience for the task product to better control for these variables.

Furthermore, students repeated four times the different tasks they had to perform for each search task. This may have influenced search behavior and outcome as students most likely started to anticipate that they would be asked to recall information source (Schoor et al., 2023). Regarding the task model measurement assignment, two limitations need to be discussed. First, we used two open-ended questions to capture students' initial task model without influencing them, as it could be the case with single choice questions. However, our measure may have underestimated the quality of students' initial task model, as participants may have refrained or failed to verbalize significant aspects of their task understanding. Despite this limitation, the measure still provides valuable insights, as most participants answered and provided at least one sentence. Furthermore, literature on individuals' ability to verbalize procedural skills also

emphasized that it can be complex to verbalize spontaneously the different actions and skills one will use during a subsequent task (Dominowsky, 1988). To solve this issue, future studies could ask individuals to rate to what extent different actions or strategies could be relevant to solve the task at hand. However, such method could also suggest some strategies and influence subsequent search behaviors (Schoor et al., 2021,2023). Secondly, participants were asked to explain their representation of the task demands prior to solving the tasks, which instructions may have increased academic expectations regarding online search behavior. Such limitations align with the self-explanation effect (Bisra et al., 2018). Indeed, asking participants to formulate their initial task model may have an influence on subsequent behavior, but this influence is not systematic. For instance, assessing students' initial task model before completing a task-oriented reading activity can support search outcome for children (Ayroles et al., 2021) or low-comprehenders adults, provided that their initial task model is relevant (Cerdán et al., 2013; Llorens & Cerdán, 2012). Some studies assessed the task model after the task completion (Schoor et al., 2021, 2023), notably to avoid the influence of the assessment of the initial task model. However, as context is postulated to influence the elaboration of the task model at the beginning of the task, we were mostly interested in examining the students' initial task model. Future works should thus examine how to assess the initial task model without being too invasive or without providing too many hints, at the risk of influencing students information processing and navigation.

Finally, it cannot be ruled out that the context in which the students were when taking this experiment also represents a limit. Indeed, students took the experiment in their university, in a classroom setting, which may further increase academic-related expectations in terms of norms. Recent studies highlighted that information searching can be performed with a wide variety of tools (smartphones, tablets) in very different settings (Kammerer et al., 2018). Future studies could explore how other features of the task context, such as the physical environment (e.g. at home, outside while waiting for the bus) or the devices used (e.g. a personal smartphone, impact task processing (Kammerer et al., 2018; McLaren et al., 2022).

Perspectives

Should the task model be measured after offering students the possibility to interact with the documents or/and the search engine? In line with the RESOLV model and recent works, a *minima* task expectancy could be necessary to foster individuals to allocate resources to the elaboration of the task model (Britt et al., 2018; Naumann, 2019; Rouet et al., 2017). Future works could thus more particularly investigate how students perceive the demands of the task in terms of self-regulation expectancies. Next, students' willingness to allocate their attention to the elaboration of a relevant task model is probably related to their motivation (Naumann, 2019). Future studies could also examine how motivation influences students' initial task model quality. Eventually, as information searching and purposeful-reading behavior is continuously affected by students' goal (Hautala et al., 2022), studies could also examine how students' task model evolve during the task completion.

Conclusions

To the best of our knowledge, this study is the first to examine students' initial task model quality and the interplay between task model, task context and task complexity in information search with a search engine. Supporting students' ability to evaluate information search or task-oriented reading requirements is a major issue (Macedo-Rouet et al., 2024). To

develop instruction tools that support autonomous and reliable processing of information on the Internet, research needs to understand how students can use clues from the task context and task instructions to engage in adapted search strategies. An academic-related context can support the elaboration of an initial task model that contains higher standards of coherence, which can foster deeper navigation for simple search tasks and deeper information and source processing for complex tasks. Findings highlighted that: (i) students' understanding of the context can influence their initial task model, and (ii) students initial task model can influence search behavior and strategies. These empirical data can bridge two fields of research: purposeful reading with multiple documents and information searching with a search engine. The theoretical considerations of the RESOLV model (Britt et al., 2018; Rouet et al., 2017) can apply to navigation and information processing behavior with a search engine. Examining students' verbalizations of their representation of the task demands can have some limitations (Bisra et al., 2018), but the TASC framework can be useful for future studies interested in task models (McCardle et al., 2017). Educators should keep in mind that framing an information search task in an academic or a non-academic context can influence students' goal, search behavior, and most likely, what students may learn from the task. As helping students use relevant cues in the task context and task instructions can help them engage in adapted search strategies and apply the expected academic norms, teachers should help students develop transversal task-modelling skills.

Declarations

Data will be made available upon request.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Ayroles, J., Potocki, A., Ros, C., Cerdán, R., Britt, M. A., & Rouet, J. F. (2021). Do you know what you are reading for? Exploring the effects of a task model enhancement on fifth graders' purposeful reading. *Journal of Research in Reading*, 44(4), 837-858. <https://doi.org/10.1111/1467-9817.12374>
- Bisra, K., Liu, Q., Nesbit, J.C. *et al.* Inducing Self-Explanation: a Meta-Analysis. *Educ Psychol Rev* 30, 703–725 (2018). <https://doi.org/10.1007/s10648-018-9434-x>
- Bohn-Gettler, C. M., & Kendeou, P. (2014). The interplay of reader goals, working memory, and text structure during reading. *Contemporary Educational Psychology*, 39(3), 206–219. <https://doi.org/10.1016/j.cedpsych.2014.05.003>
- Brand-Gruwel, S., Kammerer, Y., Van Meeuwen, L., & Van Gog, T. (2017). Source evaluation of domain experts and novices during Web search. *Journal of computer assisted learning*, 33(3), 234-251. <https://doi.org/10.1111/jcal.12162>
- Brand-Gruwel, S., Wopereis, I., & Vermetten, Y. (2005). Information problem solving by experts and novices: Analysis of a complex cognitive skill. *Computers in human behavior*, 21(3), 487-508. <https://doi.org/10.1016/j.chb.2004.10.005>
- Britt, M. A., Rouet, J. F., & Durik, A. (2018). Representations and processes in multiple source use. In *Handbook of multiple source use* (pp. 17-33). Routledge. <https://doi.org/10.4324/9781315627496>
- Byström, K., & Järvelin, K. (1995). Task complexity affects information seeking and use. *Information processing & management*, 31(2), 191-213. [https://doi.org/10.1016/0306-4573\(95\)80035-R](https://doi.org/10.1016/0306-4573(95)80035-R)
- Campbell, D. J. (1988). Task complexity: A review and analysis. *Academy of management review*, 13(1), 40-52. <https://doi.org/10.5465/amr.1988.4306775>
- Cerdán, R., Gilabert, R., & Vidal-Abarca, E. (2013). Self-generated explanations on the question demands are not always helpful. *The Spanish journal of psychology*, 16, E26. <https://doi.org/10.1017/sjp.2013.45>
- Cerdán, R., Pérez, A., Vidal-Abarca, E. *et al.* To answer questions from text, one has to understand what the question is asking: differential effects of question aids as a function of comprehension skill. *Reading & Writing*, 32, 2111–2124 (2019). <https://doi.org/10.1007/s11145-019-09943-w>
- Dominowski, R. L. (1998). Verbalization and problem solving. In *Metacognition in educational theory and practice* (pp. 25-45). Routledge. <https://doi.org/10.4324/9781410602350>
- Dosso, C., Tamine, L., Paubel, P. V., & Chevalier, A. (2023). Navigational and thematic exploration–exploitation trade-offs during web search: effects of prior domain knowledge, search contexts and strategies on search outcome. *Behaviour & Information Technology*, 1-27. <https://doi.org/10.1080/0144929X.2023.2242514>
- Goldhammer, F., Naumann, J., Stelter, A., Tóth, K., Rölke, H., & Klieme, E. (2014). The time on task effect in reading and problem solving is moderated by task difficulty and skill: Insights from a computer-based large-scale assessment. *Journal of Educational Psychology*, 106(3), 608–626. <https://doi.org/10.1037/a0034716>
- Hahnel, C., Goldhammer, F., Kröhne, U., & Naumann, J. (2018). The role of reading skills in the evaluation of online information gathered from search engine environments. *Computers in Human Behavior*, 78, 223-234. <https://doi.org/10.1016/j.chb.2017.10.004>
- Hautala, J., Salmerón, L., Tolvanen, A. *et al.* Task-oriented reading efficiency: interplay of general cognitive ability, task demands, strategies and reading fluency. *Reading & Writing*, 35, 1787–1813 (2022). <https://doi.org/10.1007/s11145-022-10265-7>

- Kammerer, Y., Brand-Gruwel, S., & Jarodzka, H. (2018). The future of learning by searching the web: Mobile, social, and multimodal. *Frontline Learning Research*, 6. <https://doi.org/10.14786/flr.v6i2.343>
- Kammerer, Y., & Gerjets, P. (2017). "The role of search result position and source trustworthiness in the selection of web search results when using a list or a grid interface": Corrigendum. *International Journal of Human-Computer Interaction*, 33(7), 607. <https://doi.org/10.1080/10447318.2017.1304121>
- Kannianen, L., Kiili, C., Tolvanen, A., Aro, M., & Leppänen, P. H. (2019). Literacy skills and online research and comprehension: Struggling readers face difficulties online. *Reading and Writing*, 32(9), 2201-2222. <https://doi.org/10.1007/s11145-019-09944-9>
- Latini, N., Bråten, I., Anmarkrud, Ø., & Salmerón, L. (2019). Investigating effects of reading medium and reading purpose on behavioral engagement and textual integration in a multiple text context. *Contemporary Educational Psychology*, 59, 101797. <https://doi.org/10.1016/j.cedpsych.2019.101797>
- Llorens, A. C., & Cerdán, R. (2012). Assessing the Comprehension of Questions in Task-Oriented-Reading//Evaluación de la comprensión de las preguntas en lectura-orientada-a-tareas. *Revista de Psicodidáctica*, 17(2). DOI : 10.1387/Rev.Psicodidact.4496
- Macedo-Rouet, M., Saux, G., Potocki, A. *et al.* Fostering university students' online reading: effects of teacher-led strategy training embedded in a digital literacy course. *Instr Sci* 52, 1021–1054 (2024). <https://doi.org/10.1007/s11251-024-09676-6>
- Mason, L., Tarchi, C., Ronconi, A. *et al.* Do medium and Context Matter when learning from multiple complementary Digital texts and videos?. *Instr Sci* 50, 653–679 (2022). <https://doi.org/10.1007/s11251-022-09591-8>
- McCardle, L., Webster, E. A., Haffey, A., & Hadwin, A. F. (2017). Examining students' self-set goals for self-regulated learning: Goal properties and patterns. *Studies in Higher Education*, 42(11), 2153-2169. <https://doi.org/10.1080/03075079.2015.1135117>
- McCrudden, M.T. & Schraw, G. (2007). Relevance and goal-focusing in text processing. *Educational Psychology Review*, 19, 113–139 <https://doi.org/10.1007/s10648-006-9010-7>
- McLaren, B., Richey, M., Nguyen, J. E., & X., Hou, (2022). How instructional context can impact learning with educational technology: Lessons from a study with a digital learning game. *Computers & Education*, 178, 104366. <https://doi.org/10.1016/j.compedu.2021.104366>
- Naumann, J. (2019). The skilled, the knowledgeable, and the motivated: Investigating the strategic allocation of time on task in a computer-based assessment. *Frontiers in Psychology*, 10, 1429. DOI : 10.3389/fpsyg.2019.01429
-
- Powell, M. J. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives. *Cambridge NA Report NA2009/06*, University of Cambridge, Cambridge, 26, 26-46.
- Rodicio, H. G. (2015). Students' evaluation strategies in a Web research task: Are they sensitive to relevance and reliability?. *Journal of Computing in Higher Education*, 27, 134-157. <https://doi.org/10.1007/s12528-015-9098-1>
- Rouet, J. F., Britt, M. A., & Durik, A. M. (2017). RESOLV: Readers' representation of reading contexts and tasks. *Educational Psychologist*, 52(3), 200-215. <https://doi.org/10.1080/00461520.2017.1329015>
- Sanchiz, M., Amadiou, F., & Chevalier, A. (2020). An evolving perspective to capture individual differences related to fluid and crystallized abilities in information searching with a search engine. In W.T. Fu, & H. van Oostendorp (Eds.), *Understanding and Improving Information Search* (pp. 71-96). Springer. https://doi.org/10.1007/978-3-030-38825-6_5

- Sanchiz, M., Chevalier, A., & Amadiou, F. (2017). How do older and young adults start searching for information? Impact of age, domain knowledge and problem complexity on the different steps of information searching. *Computers in Human Behavior*, 72, 67–78. <https://doi.org/10.1016/j.chb.2017.02.038>.
- Sanchiz, M., Chin, J., Chevalier, A., Fu, W. T., Amadiou, F., & He, J. (2017). Searching for information on the Web: Impact of cognitive aging, prior domain knowledge and complexity of the search problems. *Information Processing & Management*, 53(1), 281–294. <https://doi.org/10.1016/j.ipm.2016.09.003>.
- Schoor, C., Rouet, J. F., Artelt, C., Mahlow, N., Hahnel, C., Kroehne, U., & Goldhammer, F. (2021). Readers' perceived task demands and their relation to multiple document comprehension strategies and outcome. *Learning and Individual Differences*, 88, 102018. <https://doi.org/10.1016/j.lindif.2021.102018>
- Schoor, C., Rouet, J. F., & Britt, M. A. (2023). Effects of context and discrepancy when reading multiple documents. *Reading and Writing*, 36(5), 1111-1143. <https://doi.org/10.1007/s11145-022-10321-2>
- Sharit, J., Hernández, M. A., Czaja, S. J., & Pirolli, P. (2008). Investigating the roles of knowledge and cognitive abilities in older adult information seeking on the web. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 15(1), 1-25. <https://doi.org/10.1145/1352782.1352785>
- Vakkari, P. (2003). Task-based information searching. *Annual Review of Information Science and Technology (ARIST)*, 37, 413-64. <https://doi.org/10.1002/aris.1440370110>
- Van den Broek, P., Lorch, R.F., Linderholm, T. *et al.* The effects of readers' goals on inference generation and memory for texts. *Memory & Cognition* 29, 1081–1087 (2001). <https://doi.org/10.3758/BF03206376>
- Von Hoyer, J., Hoppe, A., Kammerer, Y., Otto, C., Pardi, G., Rokicki, M., ... & Holtz, P. (2022). The search as learning spaceship: Toward a comprehensive model of psychological and technological facets of search as learning. *Frontiers in Psychology*, 13, 827748. <https://doi.org/10.3389/fpsyg.2022.827748>

Appendix

Appendix 1. Summary of the impact of task context and complexity on search outcome and behavior (results of mixed repeated ANOVAs).

	task context	Task complexity	task context* task complexity
Search outcome	$F(1,111) = .03, p < .87, \eta^2_p < .001$	$F(1,111) = 73.45, p < .001, \eta^2_p = .40$, simple : $M = 91.7 SE = 2.50$ and complex $M = 56.2 SE = 3.36$.	$F(1,111) = 1.64, p < .20, \eta^2_p = .015$.
Proportion of time on the search engine results pages	$F(1,111) = 13.27, p < .001, \eta^2_p = .119$	$F(1,111) = 11.08, p = .02, \eta^2_p = .102$, simple : $M = 64.1 SE = 3.717$ and complex $M = 32.4 SE = 2.91$	$F(1,111) = 4.44, p < .04, \eta^2_p = .04$. Bonferroni's comparison showed that for simple tasks, the impact of context was significant ($t(98) = 3.51, p = .004$, nonacademic $M = 57.8 SE = 5.32$ and academic $M = 34.4 SE = 4.64$), whereas for complex ones, the impact of context was not significant ($t(98) = .70 p = 1.00$, nonacademic $M = 34.3 SE = 4.18$, academic $M = 30.6 SE = 3.62$).
Number of webpages consulted	$F(1,111) = 4.63, p = .03, \eta^2_p = .04$	$F(1,111) = 32.50, p < .001, \eta^2_p = .23$, simple : $M = 1.02 SE = .11$ and complex $M = 2.08 SE = .19$.	$F(1,111) = .29, p = .59, \eta^2_p = .002$.
Source recall	$F(1,111) = 17.19, p < .001, \eta^2_p < .14$	$F(1,111) = 17.90, p < .001, \eta^2_p = .14$, simple : $M = 1.62 SE = .10$ and complex $M = 1.22 SE = .12$.	$F(1,111) = .29, p = .59, \eta^2_p = .002$

Appendix 2. means and SD for the time spent of the search engine results pages and the time spent on webpages.

		Non academic related topic	Academic related topic
Time spent on the search engine results pages	Simple	144.5 (15.0)	120.0 (15.3)
	Complex	180.2 (21.7)	89.3 (14.6)
Time spent on websites	Simple	125 (16.1)	188 (18.0)
	Complex	424 (28.7)	447 (30.4)