Didrik Telle-Wernersen

Perceived usefulness of Boolean queries

A user-centered study

Master thesis 2021
Master in Library and Information Science

OsloMet – Oslo Metropolitan University,
Department of Archivistics, Library and Information Sciences

Abstract

This master thesis examines how users in an information retrieval situation perceive the usefulness of simple and advanced Boolean queries. Graduate students from an educational science program at a university in Norway were recruited to take part in the project. Participants were in the beginning of writing their master thesis at the time of data collection. Data collection methods consisted of pre-search interviews regarding the particpants' master thesis topic and information needs, think-aloud interviews focusing on relevance assessements of retrieved results, and post-search interviews concerning their perceived usefulness and experiences with using Boolean queries when searcing. The results of one simple and one advanced Boolean query were assessed by the participants, who stated their relevance criteria and relevance assessments while interacting with the ERIC database via the EBSCOhost interface. Data collection was done digitally using the meeting service Zoom. A total of 7 participants assessed 149 results verbally. Results from a directed content analysis of the interview and interaction data indicates that percevied usefulness of Boolean queries should be understood in the light of information need types, and that situational relevance can inform this understanding.

OsloMet – Oslo Metropolitan University,
Department of Archivistics, Library and Information Sciences

Preface

This master thesis has been a very interesting experience, and at times quite a demanding one. Sincerly thanks to my supervisor professor Pia Borlund for all the thought-provoking conversations along the way, and for help in guiding the project to the finish line. Thanks also to the participants, for sharing their time.

I also want to express my gratitude towards my colleagues for all their assistance, and to my friends and family for all their help. Finally, my heartfelt thanks to May and Mads, for all your patience and support!

Didrik Telle-Wernersen

December 2021

Contents

List of figures

List of tables

1 Introduction

Teaching students how to use Boolean operators and search techniques can be said to be an integral part of academic library instruction sessions within higher education in Norway, both at undergraduate and graduate levels. Information on how Boolean search works is presented on Norwegian academic library websites[1,2], on web resources such as Search & Write[3], and included in recent course specific literature in higher education, e.g. in Høgheim (2020, Chapter 3)**.** Instruction on Boolean retrieval and database searching is also still considered an important part of Library and Information Science (LIS) training (Hjørland, 2015, p. 1562).

Empirical research does however indicate that college students experience difficulties both with finding appropriate terms that describe their information needs, and with combining such terms with Boolean operators when searching (Dempsey & Valenti, 2016, p. 205). Research has also shown that few graduate students utilize advanced search techniques, including Boolean operators (Catalano, 2013, p. 267). At the same time, searching via Google and the introduction and increasing use of library discovery services, both with a single search box and simple keyword search as default, have challenged the way both librarians and students think about search and research practices (Georgas, 2013, p. 166; Hamlett & Georgas, 2019, p. 231).

Boolean operators are typically used with the intent to control the precision of a query, and to gain comprehensiveness by including synonyms and related terms. This is especially important in systematic searches used when gathering literature for a systematic review or meta-analysis (Gusenbauer & Haddaway, 2019, p. 182). The need for comprehensiveness in a literature search for a student assignment will necessarily vary, according to level of education and field of study, and across research traditions. The evidence-based practice paradigm within health science places great emphasis on a systematic approach to searching to gain comprehensiveness, also at an undergraduate level in Norwegian higher education (Strømme, 2019). Characteristics of this approach include the employment of PICO (population, intervention(s), comparator(s) and outcomes) schemes, a controlled vocabulary

[1] https://innsida.ntnu.no/wiki/-/wiki/English/Advanced+literature+search
[2] https://film.oslomet.no/soketeknikk-boolske-sok-med-and-og
[3] https://sokogskriv.no/en/searching/searching-techniques.html#combine-search-terms

of subject terms, and the use of Boolean operators to combine, include and exclude controlled and uncontrolled terms (Lefebvre et al., 2020).

This master thesis aims to explore how graduate students within another field of study, the educational sciences, perceive the usefulness of advanced Boolean search strategies when searching subject databases. The introduction in 2017 of a compulsory five-year graduate degree for all new primary school teachers in Norway requires more specialized knowledge and greater ability to utilize research-based teaching methods among graduated candidates than earlier (Ministry of Education and Research, 2016). The author's view is that this context might place a greater demand on educational science students' search skills. Further, it might generate more complex search tasks and information needs among these students, and possibly increase the need for comprehensiveness in literature searches. An exploration of what kind of experiences and opinions students within the educational sciences have with using simple and advanced Boolean search strategies in this context, is therefore of interest.

In two recent studies, Lowe et al. (2018; 2020) compared the relevance of results across databases and discovery services when using queries containing natural language as opposed to simple Boolean queries, and simple Boolean as opposed to advanced Boolean queries. Results from both studies indicate that there are small differences in the performance of these three different search techniques (natural language, simple Boolean, and advanced Boolean), in terms of topical relevance. These findings suggest that new considerations could be made regarding the scope of academic library instruction for students, and how this instruction reflects students' actual information needs, and subsequent search habits.

The research questions in this thesis are inspired in particular by Lowe et al. (2020), who in their study explored the potential benefit of using the Boolean operators 'OR' and 'AND' in queries as opposed to only the Boolean operator 'AND', for students who are searching for literature for their master's thesis or similar projects. In their study, topical relevance is used as a performance measure without the involvement of actual users and their information needs and relevance judgements. How relevance is understood, defined, and used as a performance measure in information retrieval (IR) research is however subject to continuous investigation and debate (Saracevic, 2017). This made the author curious to see if a user-

centered study on the same research themes as those explored by Lowe et al. (2018; 2020) could be achieved, and what the results would indicate, in terms of potential benefits of using Boolean operators for these users.

The first part of the title of this thesis is inspired by the concept of 'Perceived usefulness', defined by Davis (1989, p. 320) as: "the degree to which a person believes that using a particular system would enhance his or her job performance". This can be contrasted with the concept of 'Perceived ease of use', defined as "the degree to which a person believes that using a particular system would be free of effort" (Davis, 1989, p. 320). The 'particular system' in this context is the Boolean operators 'AND' and 'OR'. 'Job performance' is the achievement of a work task. The theoretical perspective in this thesis does however not involve the related Technology Acceptance Model (TAM) proposed by Davis. The focus is instead directed at how the concept of relevance and the concept of information needs in IR can be used to understand the perceived usefulness of Boolean queries.

This thesis explores the perceived usefulness of simple and advanced Boolean queries as viewed by the users themselves. By employing methods from the interactive information retrieval (IIR) field of research and including authentic information needs and relevance assessments made by students, insights into the opinions and experiences of the user group in question are provided. Hence, this master thesis produces new empirical data on the perceived usefulness of Boolean search strategies in a real context. This in turn, can help inform pedagogical approaches for academic librarians and the design of instruction material in classes and on the web. Consequently, the focus is not on gaining insight which can directly aid in the development of IR systems as such, but rather on how users perceive the utility of certain retrieval techniques in such a system. The research questions are as follows:

> *RQ1: How do graduate students within the educational sciences perceive the usefulness of simple and advanced Boolean queries, when searching a subject database?*
>
> *RQ2: How is the perceived usefulness of simple and advanced Boolean queries among students related to the students' information needs and relevance criteria?*

To provide data that would aid in answering these questions, educational science graduate students in the beginning of their master thesis projects were recruited as participants. They were asked to take part in mediated search sessions where one simple and one advanced

Boolean query were used. The queries were constructed in collaboration with each participant, using their own reported information need as the starting point. Searches were carried out in the ERIC (The Education Resource Information Center) database, via the EBSCO*host* interface. Search terms used in the queries were reviewed by the author and each participant in collaboration, and the participants were asked to comment on the appropriateness of each query with regards to their personal information need. This was done with the aim of reflecting the underlying and potentially dynamic information need to the greatest extent possible.

The participating students were asked to assess the relevance of retrieved results for each of the two queries, and to state these assessments in a think-aloud manner. Verbalizations of relevance criteria were then analyzed for frequency to give insight into the relevance assessment process. Pre-search interviews were conducted to provide context with regards to the participants' information needs, the current stage of work task process, familiarity with their chosen master thesis research theme and intended use of found information. Post-search interviews focused on participants' reflections on the perceived usefulness, quality and complexity of the queries used.

2 Theoretical perspective

To establish a foundation for the theoretical perspective on IR research in this thesis, Wilson's (1999) nested model (Figure 1) is useful as a starting point. In his paper, Wilson (1999) points out that research efforts and utilized models concerning information behavior describe phenomena and processes that exist at different levels. Placed at the outmost layer in this model, the research area of information behavior focuses on the macro level, examining broad questions encompassing interaction with, use of, and search for information. The second layer concerns the discovery of and access to information sources (information-seeking behavior research). At the inner layer, the micro level, the mediated and un-mediated interactions between users and systems are investigated (information search behavior research). These layers can be thought of as divisions of the broader field of information behavior research (Wilson, 1999, p. 263).

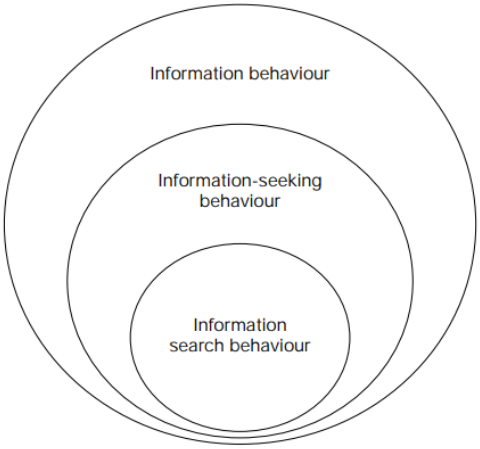


Figure 1: "A nested model of the information seeking and information searching research areas" (Wilson, 1999, p. 263).

As this thesis is concerned with the interactions between user, mediator, and system, and more specifically the perceived usefulness (among users) of certain retrieval techniques employed in such interactions, it is placed in the innermost research field of information search behavior. The inner field of Wilson's model is understood as including IIR research (Ingwersen & Järvelin, 2005, p. 198; White & Roth, 2009, p. 37).

2.1 The cognitive viewpoint in IR research

Wilson's model depicts research fields at different levels which all concentrate on behavioral, and hence, user-oriented phenomena (White & Roth, 2009, p. 37). In research that deals with user-system interactions and evaluations of the performance of IR systems, a system-oriented perspective is also present. This research tradition and its resulting evaluation exercises, including the Text REtrieval Conference (TREC), can be traced back to the IR system experiments known as the Cranfield tests that were carried out in the late 1950's to the mid 1960's (Kelly, 2009, p. 13; Robertson, 2008, p. 439). Crucial to the system-oriented evaluation tradition is the use of test collections on which different IR systems are tested for effectiveness (Clough & Sanderson, 2013).

In addition, relevance is often based on topicality, or the 'aboutness' of a retrieved result, and treated in a binary way (a retrieved result is judged as either relevant or not-relevant), enabling calculations of effectiveness measures such as recall and precision (Sanderson, 2010). The notion of binary and non-binary relevance, and different types of relevance is elaborated and discussed in section 2.3. Some of the elements in the system-oriented

perspective are discussed further in the subsequent sections, but the tradition as such is not treated at length here. For further information on the system-oriented tradition, the reader is referred to the overview and discussion by Sanderson (2010).

In the inner field depicted in Wilson's model, user-system interactions have been conceptualized in numerous ways resulting in models that focus on various aspects, depending on the aim of the research. One such example is the cognitive model of IR interaction (Figure 2) developed in Ingwersen (1996) and extended in Ingwersen (2000, p. 159). It is chosen here to illustrate the components involved in interactions that take place between the user, the system, and the intermediary during an IR situation. In this model, all components are thought of as representing cognitive structures: The user is represented by his or her cognitive space, containing e.g., the information needs, perceptions of the current work tasks and search tasks, and all previous experiences and resulting expectations of the IR situation. All these factors (and in essence, an unknown number of others) within a user constitute the individual user side of the interaction. The social and organizational environment seen on the right side of the model provides the context in which the user's information search takes place.

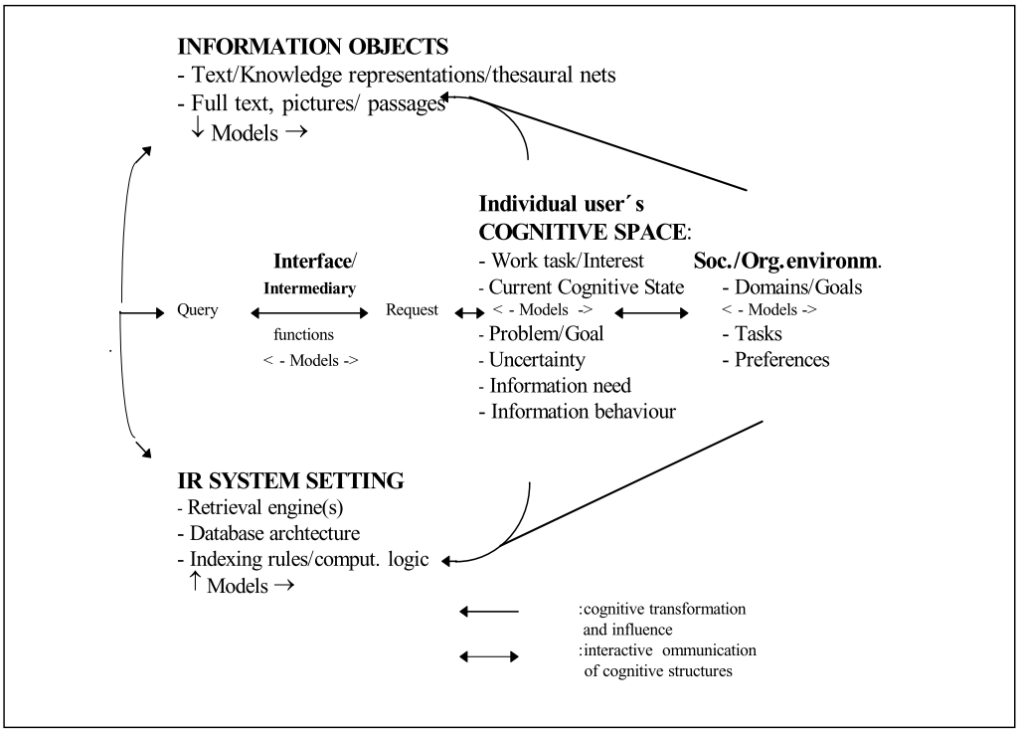


Figure 2: Cognitive model of IR interaction. From Ingwersen (2000, p. 159)

On the IR system side to the left, cognitive structures are present in the system design, knowledge organization choices and principles. These are expressed through thesauri containing controlled terms, and assigned subject terms, as well as the retrieval and ranking algorithms processing the queries and information received by the system. Another set of such structures are present through the documents retrieved in the system. The intermediary or interface seen in the center of the model acts as yet another cognitive structure, processing information requests that in turn are formulated as queries acceptable to the system.

Ingwersen's cognitive IR model is chosen as a frame of understanding due to its focus on the user-intermediary interaction, as pointed out by Savolainen (2018, p. 978). Sormunen (2000, pp. 31-32) identifies the interface/intermediary component as cognitive structures that are engaged in query formulation. The intermediary role as engaged in query formulation and user interaction in the resulting research design will be discussed in section 4.1. The idea of a system of interacting human cognitive structures is fundamental in understanding what is known as the cognitive viewpoint of IR: That changes occur in these structures during IR interaction, resulting in information needs, relevance assessments and tasks that are dynamic and situational in nature (Ingwersen, 1999, p. 3)

The theoretical, and in turn methodological perspective in this thesis resides within this viewpoint. Kelly (2009, p. 16), drawing on the work of Ingwersen (1996) states that the viewpoint "embraces the complexity inherent in IR when users are involved and focuses attention on the cognitive activities that take place during information seeking and retrieval, and user-information, user-system interactions". Research efforts within this viewpoint then, include test participants with real information needs, and the observation of relevance assessments made by the participants themselves. Borlund (2010, p. 24) underlines the fundamental role of the test participant's own information need(s) within this perspective, in that the cognitive viewpoint "is about the user's desire for information, and hence is a platform for authentic information studies of users' retrieval, search and seeking interactions in the process of achieving this goal of desired information".

In the research design developed for the present thesis, the participants' 'goal of desired information' is intended to play a key role. The focus is on how users perceive the usefulness of certain retrieval techniques while interacting with an IR system in the context of an

individual and dynamic information need. The theoretical concept of information need will be discussed in section 2.4.

2.2 Realism and control

Kelly (2009, pp. 16-17) gives an interesting consideration of whether IIR research is primarily done within a system- or a cognitive perspective. The main challenge in this respect is that of balancing experimental control on one hand, with realism on the other. As a result of individual and unobservable cognitive processes that are ongoing throughout a user's interaction with the IR system, as indicated in Ingwersen's model (fig. 2), each user will interact differently with the system (Kelly, 2009, p. 16). While this user activity provides realism in an IIR study, it decreases the possibility of creating experimental conditions that allow for replications and generalizations. This challenge is underlined by Belkin (2008, p. 52), who highlights the challenge as a need for realism, and a desire for generalizations. This challenge in balancing realism and control is revisited in section 4 on methodology. In the following sections, theoretical concepts considered crucial to IIR research efforts are presented and discussed.

2.3 Relevance in IR research

Relevance can be characterized as the basic objective and intent of all IR systems (Borlund, 2003, p. 913; Saracevic, 2017, p. 1). In other words, the design and development of such systems can be understood as governed by the overarching goal of providing users with relevant information. Since the first IR systems were built and tested in the 1950's, information science as a research field has dealt with the concept of relevance in a number of ways, and it has been subject to much deliberation. This is illustrated in a historical overview by Mizzaro (1997), where relevance research from 1958 until 1997 is surveyed. In the reviewed literature a myriad of definitions, measurements and understandings of the concept are present through a diverse set of conceptual, experimental, and theoretical papers. However, in brief, the overarching trend in the literature observed by Mizzaro is characterized as a shift from relevance seen as an underlying property of documents and queries (although a user-based view is also present), towards research that directs its efforts at examining and understanding the nature of a "subjective, dynamic and multidimensional relevance" (Mizzaro, 1997, p. 827).

This changing focus is represented in a central paper by Schamber et al. (1990) who conclude their survey of the relevance literature by proposing an alternative view on the concept. This alternative revolves around the user as the central factor in determining relevance, and is formulated through three conclusions:

1. Relevance is a multidimensional cognitive concept whose meaning is largely dependent on users' perceptions of information and their own information situations.
2. Relevance is a dynamic concept that depends on users' judgments of the quality of the relationship between information and information need at a certain point in time.
3. Relevance is a complex but systematic and measurable concept if approached conceptually and operationally from the user's perspective. (Schamber et al., 1990, p. 774)

These conclusions set the stage for subsequent research efforts in the user-oriented tradition and emphasize the role of the user's information need. In the context of evaluation of IR systems, Robertson and Hancock-Beaulieu (1992) identified three revolutions, in which the 'relevance revolution' relied on the notion of relevance assessed in relation to a user's need rather than the request submitted to an IR system.

As a result of the two other revolutions pointed out by Robertson and Hancock-Beaulieu (1992), the cognitive and interaction revolutions, topical relevance was no longer the only performance measure in IR experiments, but came to be considered alongside multi-dimensional and dynamic relevance assessments belonging to an individual user (Ingwersen & Järvelin, 2005, p. 3). The multi-dimensional aspect is understood as relevance expressed through "classes, types, criteria, degrees, and levels of relevance" (Borlund, 2003, p. 914), and is discussed further in the sections below.

2.3.1 Dimensions of relevance

In terms of conceptual definitions of relevance within the user-oriented research tradition, Saracevic (1996b) proposed the idea of a system of distinct but also interdependent types of relevance, which came to be influential in subsequent IR research. This system is expressed through five manifestations of relevance, based on the following relations:

> *System or algorithmic relevance:* relation between a query and information objects (texts) in the file of a system as retrieved, or as failed to be retrieved, by a given procedure or algorithm. [...]
>
> *Topical or subject relevance:* relation between the subject or topic expressed in a query, and topic or subject covered by retrieved texts, or more broadly, by texts in the systems file, or even in existence. [...]
>
> *Cognitive relevance or pertinence:* relation between the state of knowledge and cognitive information need of a user, and texts retrieved, or in the file of a system, or even in existence. [...]
>
> *Situational relevance or utility:* relation between the situation, task, or problem at hand, and texts retrieved by a systems [*sic*] or in the file of a system, or even in existence. [...]
>
> *Motivational or affective relevance:* relation between the intents, goals, and motivations of a user, and texts retrieved by a system or in the file of a system, or even in existence. [...] (Saracevic, 1996b, section 4, (italics in the original))

Harter (1992) defined two main classes of relevance: One that is objective, or system-based, and a subjective, or user-based (p. 602). The five relevance types proposed by Saracevic (1996b) are by Borlund (2003, p. 914) defined as belonging to these two classes in the following way: System or algorithmic relevance represents objective relevance, and the remain four reside in the subjective class. The two classes represent perspectives on

relevance in IR research and development of IR systems, and can further be said to parallel the system-driven and the user-oriented research traditions, respectively (Borlund, 2003, p. 914).

Hjørland (2010, p. 218) challenged the concept of an 'objective' class of relevance, as relevance assigned by a system is the result of decisions that were originally human (choice of input in retrieval algorithms, ranking functions etc.) and therefore subjective in its origin. A similar analysis is put forward by Saracevic (2017, pp. 24-25). This argument is still in line with the cognitive viewpoint presented in section 2.1., in which the IR system and its architecture is considered exactly this: The representations of cognitive, and hence, human knowledge structures (and initially, decisions based on relevance assessments) expressed through e.g., retrieval algorithms. The class of objective types of relevance is understood here as it is described by Borlund (2003, pp. 914-915): It is represented by algorithmic relevance assigned to results by the IR system, based on topicality as expressed in the contents of the retrieved objects.

Common for all the remaining types of relevance, those that belong to the subjective class, is that they all involve "various *degrees of intellectual interpretations* carried out by human observers—whether assessors or users" (Borlund, 2003, p. 915, italics in the original). In other words, relevance belonging to this class is always based on a user's subjective opinion. As seen above, topical relevance is by Saracevic (1996b) defined as based on a relation between topic expressed in a query and topic covered by retrieved information objects. When this topicality is assessed by a user, whether the assessment is intended as subjective or objective, it is known as 'intellectual topicality' (Borlund, 2003, p. 915). It is influenced by the cognitive structures present in the user.

The cognitive relevance type, also called pertinence, is manifested through a relation between the user's information need and the retrieved objects, and through the need, it is linked to a user's knowledge. Borlund (2003, p. 915) states that this relevance type enables the existence of a dynamic information need, as cognitive relevance relies on interpretations based on existing knowledge, which then can be impacted and result in a change in the information need.

Situational relevance manifests itself through relation between the retrieved objects, and a user's underlying situation or task. It is connected to the usefulness of results. It is therefore considered to be a context-dependent type of relevance, but since the task is considered as the background for the existing information need in a user, it can be challenging to distinguish from cognitive relevance (Borlund, 2003, p. 915) Recent conceptual developments of situational relevance include work by Jiang et al. (2017) and what they term Ephemeral State of Relevance, where the focal point is the amount of useful information a user is able to obtain from each result in the moment of assessment, narrowing the scope of analysis.

Motivational relevance is the fifth relevance type proposed in Saracevic's system. It is by Cosijn and Ingwersen (2000) and Borlund (2003, p. 915) not considered as a separate type of relevance, but rather the underlying drive or motivation that affects the other types of subjective relevance. The view is shared by this author.

2.3.2 Relevance as a performance measure

One of the classic performance measures in the system-oriented IR tradition is that of recall and precision, which is derived by establishing how many of the relevant objects in a collection that were retrieved (recall), and how many of the retrieved objects that are relevant (precision) (Kelly, 2009, p. 109). Calculations of this measure depends on binary relevance, which rely on assessment of topical relevance. This brings about the dimension of *degrees* of relevance. Non-binary, or graded assessments of relevance is related to the cognitive IR model by Ingwersen, as they allow for the identification of partially relevant results (1996, p. 43). Development of evaluation measures based on graded assessments include efforts by Kekäläinen and Järvelin (2002), where briefly stated, the graded assessment of an object and rank in a result list informs the performance measure cumulated gain (CG) and cumulated gain with discount (CGD).

A 'collapse' into two categories of relevance is sometimes seen in research that uses graded relevance assessments, as pointed out by Spink et al. (1998, p. 600): partially relevant, and relevant are combined into 'relevant', and similarly, partially not relevant and not relevant are combined into 'not relevant'. Although this enables calculations of recall and precision measures, but the nuances of partially relevant results become hidden. The classic recall and precision measure can also be contrasted with the concept of 'perceived precision' used in a

study by Vakkari and Huuskonen (2012), where it is defined as according to the number of results the user found useful, in the sense that the results contributed to a user's task. The user's own perception is considered prominent in this regard: "The system and the search are as good as the searcher perceives them to be and cannot be disputed by anybody else. In that sense, perceived precision is a valid and reliable indicator of search output" (Vakkari & Huuskonen, 2012, p. 661).

Topical relevance continues to be used as an indicator of retrieval effectiveness, as is the case in the previously mention studies by Lowe et al. (2018; 2020), where graded relevance assessments also are collapsed into binary ones to calculate precision measures as one aspect of the result of these studies. A recent study by Vakkari et al. (2019) has sought to further complement the use of topicality assessments, by examining the user's perceived usefulness of search results as a measurement. In Vakkari et al.'s study, this is done through examining what results were used in the process of a writing task.

Measurements of the actual use of retrieved results, for example in a completed student assignment does necessarily require a longitudinal research design (Ingwersen, 2000, p. 162). Owing to limited opportunity in doing a longitudinal study in this thesis project, the scope of data collection is restricted to the assessments of relevance made by participants engaged in a session-based interaction with an IR system. One approach to examining the chosen research questions in this setting could be to identify user-based statistical relevance measures such as CGD, based on sets of retrieved results. However, during the development of this project, certain theoretical elements of the concept of relevance itself emerged as particularly interesting to consider. Inspired by previous empirical research on relevance criteria examined in section 3.3, a decision was made to look more closely at criteria and their relation to the different types of relevance proposed by Saracevic (1996b).

To examine what relevance types were involved in the data collection sessions and how they came to manifest themselves, focus was directed at the acts of relevance assessments, in the form of verbalized criteria stated by the participants. The data collection procedure and analytical approach described in section 4 therefore focus on the criteria and type dimensions of relevance. Relevance criteria are defined here as "the parameters by which the users determine the relevance of the retrieved information objects" (Borlund, 2003, p. 917). The dimension of relevance criteria is not treated at length in this section, as is it will

be revisited in the subsequent section on earlier research. Frequencies of relevance criteria that are stated by participants provide the basis for an analysis of what kind of relevance types are engaged during the data collection sessions. This in turn, is anticipated to provide clues to how usefulness of Boolean queries is perceived by participants, when seen in the context of an information need.

2.4 Information needs

Another important concept in IIR studies is that of information needs, and the perceptions of how such needs arise and how they can be characterized. One early and important contribution to this understanding was given by Taylor (1968). He describes a process in which a person's need for information develops along a continuum of stages, from being vague, inexpressible and perhaps represented by "only a vague sort of dissatisfaction" (p. 182), then gradually becoming more and more concrete. Along this continuum, he depicts four levels of a process he named 'Question formation':

> Q1—the actual, but unexpressed need for information (the *visceral* need);
>
> Q2—the conscious, within-brain description of the need (the *conscious* need);
>
> Q3—the formal statement of the need (the *formalized* need);
>
> Q4—the question as presented to the information system (the *compromised* need)
>
> (Taylor, 1968, p. 182 , (Italics in the original)).

As the description of Q4 suggests, this version of an information need is the result of a negotiation taking place between the actual need experienced by a user and the limits set by the system receiving the request in the form of a query. Taylor then suggests five filters that are activated during interactions between users, and intermediaries assisting them in IR. They are:

1. determination of subject;
2. objective and motivation;
3. personal characteristics of inquirer;

4. relationship of inquiry description to file organization;

5. anticipated or acceptable answers (Taylor, 1968, p. 183).

The filters are described as not being mutually exclusive, and that several of them might be included in the same statement from a user during such interactions. These five filters inspired the pre-search interview approach described later in section 4.5, as they point out key components of the user-intermediary interaction during the pre-search interview and query formulation, which in essence leads to a Q4-need.

Taylor's perspective was on the interaction and communication aspect of the libraries of his time and the libraries' potential for innovation in this regard. It proved highly influential in the following decades of user-centered library research. In this 1968 text the treatment of IR interaction and relevance as dynamic and situational, user-centered concepts still comes across as relevant, and is therefore of great value to the theoretical perspective in this thesis.

A development of the ideas proposed by Taylor (1968) can be seen in the ASK (Anomalous state of knowledge) hypothesis formulated by Belkin et al., and most famously described in Belkin et al. (1982a). The hypothesis states that "an information need arises from a recognized anomaly in the user's state of knowledge concerning some topic or situation and that, in general, the user is unable to specify precisely what is needed to resolve that anomaly" (Belkin et al., 1982a, p. 62). Here, the IR system-user interaction might be said to be the main context of discussion. Belkin and colleagues challenge the best-match principle implemented in IR systems of the time, which they define as relying on the matching of requests for information in the form of queries and/or index terms on one hand, and textual content in the system on the other. They identify two underlying assumptions to this principle, which they consider to be unrealistic and therefore problematic: That a user is able to identify and express all facets of an information need in a request put to the system, and that there exists a functional equivalence between a user's stated information need, and textual representations in the system (Belkin et al., 1982a, p. 63).

Both assumptions relate to the concept of Taylor's Q4 need, understood as the result of a compromise. They respectively relate to the challenges users *and* their intermediaries might

meet in transforming an information need into a representative query. The resulting design study by Belkin et al. (1982b) consequently focused on developing IR systems that relied on the process of a user explaining an underlying anomalous state of knowledge, rather than to have the user state the information need as a request to an IR system (Belkin et al., 1982a, p. 62).

The introduction of the ASK hypothesis represents a significant change in the understanding of the nature of information needs, and how they are operationalized in IR research (Borlund, 2010, p. 27). Departing from the system-oriented view of information needs as a stable concept disconnected from context, the user-oriented IR research tradition embraced information needs as connected with the individual user, contextual and potentially dynamic in nature.

Building on the insight represented in the work by Belkin et al. (1982a, 1982b), Ingwersen (1996, p. 15) pointed to the creation of a user's ASK and ensuing information needs as a result of a user's tasks or interests. These tasks and interests in turn, are "strongly influenced by the domain and dominated by the individual intentionality and cognitive state" (Ingwersen, 1996, p. 15). Domain is interpreted as e.g., field of study or research tradition, and offers a contextual view of a user's information need, in line with the social and organizational elements on the right side of Ingwersen's cognitive IR model (Figure 2, section 2.1).

In relation to knowledge domain and the cognitive state of a user engaged in IR, the concept of 'Label effect' is also considered as an important phenomenon to include in a study that deals with queries based on users' needs. The effect implies that a user will apply few or even just a single term or concept to verbally describe his or her need, often as a consequence of uncertainty, or assumptions as to what is acceptable by an intermediary or an IR system (Ingwersen, 2000, p. 164). In an effort to diminish potential label effects in the data collection procedure, both written and verbal statements were collected from participants to construct queries. This strategy is elaborated in the methodology chapter, in section 4.5.2.

To connect the dynamic nature of information needs with task domain, user intentions and cognitive structures, information need types as they are described by Ingwersen (2000) are

incorporated here. In this typology shown in Figure 3, the types of needs are characterized along the horizontal dimension by the user's own perception of the need, or how well-defined it is in accordance with the user's work task, and along the vertical as to how variable the need might be over time (Ingwersen, 2000, p. 164).

Intrinsic information need vriables – given a perceived work task	Well-defined	Ill-defined
Stable	***Verificative*** ***Conscious topical*** *Querying* Filtering behaviour	***Muddled task & info.need*** Search loops
Variable	***Conscious topical*** *Query-Navigation* Dynamic interaction	***Defined work task*** ***Muddled info.need*** *Browsing* Try-&-error behaviour

Figure 3: Typology of information needs from Ingwersen (2000, p. 163)

Three kinds of needs are defined as based on the following situations for users:

Verificative information needs (VIN), in which the user "wishes to verify information objects with known non-topical (structured) data, such as author names, client address, cited authors, journal name, etc. This type is assumed to be stable during a session period until objects have (not) been retrieved" (Ingwersen, 2000, p. 164). In terms of relevance assessments in the context of such needs, a user can be expected to be capable of assessing both the topicality and cognitive relevance (pertinence) of results based on existing knowledge and certainty, situational relevance on the basis of work task perception, as well as engaging in query modification (Ingwersen, 2000, p. 165).

Conscious topical information needs (CIN), which implies "that the user wants to clarify, review or pursue information in known subject matter and domain. Known subject matter signifies topical (unstructured) data on contents, such as terms, concepts, image representation, etc." (Ingwersen, 2000, p. 164). Needs of this type are assumed as either stable (verificative) or variable within an IR session. Topicality and cognitive relevance assessments, as well as query modification, might be influenced by uncertainty in the

variable conscious topical type (Ingwersen, 2000, p. 165), although a well-defined work task enables assessment of situational relevance.

Muddled information needs (MIN), which signal that "The user is engaged in the exploration of new concepts and relations outside known subject matter or domain, or the known data are incomplete and cognitively vague" (Ingwersen, 2000, p. 164). Label effect is by Ingwersen suggested as a prominent feature of both stable and variable version of these needs, combined with a high degree of uncertainty (2000, p. 165). He further argues that the context of variable, muddled needs might make assessments of topical and cognitive relevance quite challenging for a user, however, situational relevance is suggested possible given a defined work-task. In the case of an ill-defined work task, "severe difficulty" is associated with all types of relevance assessments and query modifications (Ingwersen, 2000, p. 166).

A user may transition between information need types as a result of viewing and assessing retrieved information objects in an IR system, reflecting the understanding of information needs as potentially dynamic. The search behaviors associated with each information need type proposed by Ingwersen (2000) was later verified empirically by Borlund and Dreier (2014). Observation of search behaviors connected to the information need types as such, dictates that participants would search the web or a bibliographic database in their own manner. Although this was the ambition in the initial research design in this thesis, it is not accommodated in the revised, final version of the design. The typology is still considered very useful as a framework for understanding the context of participants needs and the relevance assessments made, and the inclusion of query modification as an activity related to these assessments.

2.4.1 The concept of task

As pointed out by Ingwersen (1996, p. 15) and embedded in the information need typology from Ingwersen (2000), the formation of a user's information need and subsequent interaction with an IR system can be viewed as a result of an ongoing task. As with relevance, task in IR research is a debated concept with varying definitions and understandings (Soufan et al., 2021). Byström and Hansen (2005) propose a framework for this concept intended for empirical studies of information behavior, which is used in this

thesis. In their 'process view' of tasks, the ongoing task is exhibited through the activities the user performs, to progress in or solve the task itself (Byström & Hansen, 2005, p. 1051).

Further, in a generalized way a task can be seen as existing in three stages: through its construction, performance, and completion (Byström & Hansen, 2005, p. 1055). In terms of understanding the levels of different tasks for users, 'work task' is seen as the overarching assignment or project the user is involved in, which then can lead to, and be divided into 'information seeking tasks', and further into 'information search tasks' or 'information retrieval tasks' (pp. 1055-1057). Search or retrieval tasks can be performed in connection with any stage of an overarching work task.

Relating to back to Ingwersen's information need typology presented in the previous section, tasks can be considered in varying degrees of complexity, affecting the nature of the generated information need. In relation to the task process, information needs are seen as situational in nature because they are the result of a perceived 'lack of information' which is then acted upon (Byström & Hansen, 2005, pp. 1054-1055).

In the terminology presented in Byström and Hansen (2005) and Ingwersen (2000), the final research design described in section 4.7 involves participants who are engaged in information retrieval based on a potentially dynamic information need of a given type. The retrieval task and resulting information need is intended to be connected to their ongoing work task of planning and writing a master's thesis, and to result in relevance assessments made in the context of individual information needs and work task situations. The work task is considered as being in the construction stage for all participants.

2.5 Query formulation and modification

Continuing in the same terminology, query formulation might be considered a sub-process taking place at the beginning of an information retrieval task. This initial process can then be seen as dependent on the dominating information need type as indicated in Ingwersen's typology, where the act of modification of an initial query is connected to relevance assessments. Wacholder (2011) emphasizes the challenges in observing these formulation processes, as they are essentially cognitive processes taking place inside the user's mind, only observable through the product (the queries submitted to the IR system) or through "other externally observable behaviors" (p. 157).

Keeping the theoretical treatments of information needs in mind, what is a Boolean query in this perspective? Using the levels proposed by Taylor (1968), a query can be understood as an expression of a compromised (Q4) need (Byström & Kumpulainen, 2020, p. 8), adapted to the syntax accepted by the IR system. This includes a compromise between the user's need and the logic behind Boolean operators, which can be especially demanding for student users who have little training in, or seldom use such techniques. Frické (2021, p. 189) suggests that challenges in using Boolean operators can arise because of the unfamiliar use of the logical conjunctions and disjunctions represented by 'AND' and 'OR'. Not only does the user face the challenge of finding the correct, and often subject-specific terms in which to express his or her need, but also the challenge of adopting a logic that might at first seem counter-intuitive to everyday language. Filter no. 4 (Taylor, 1968) 'relationship of inquiry description to file organization' concerns both of these compromises in a user-intermediary interaction, as the information need has to be reflected in the choice of each single term, and their final combination through the use of operators.

An extension of Taylor's Q4 compromise can perhaps be argued to exist when queries are to be constructed in another language than that which is native of the user. Vanopstal et al. (2012) explored the effect of language skills on information need articulation, query formulation and relevance judgement among Dutch nursing students searching the PubMed data base with queries and results in English. The model used in their study identifies translation actions taking place during the articulation of an information need and the resulting query construction, and during relevance assessments. The results do however not show any significant correlation between language skills and quality of search terms used (Vanopstal et al., 2012, p. 1551). The native/non-native language aspect of query formulation in English is an interesting research direction in its own right. It is considered here as a part of the context for the user group in question, but it is not chosen as a main focus for the resulting research design.

Hjørland (2015) highlights the value of Boolean systems and search in "[…]providing users with the power to make informed searches and have full control over what is found and what is not" (p. 1559). This control-perspective is interesting, as it emphasizes the potential personal gain for users in employing such techniques. It might also be interpreted as the underlying perspective in academic library instruction focusing on information literacy skills.

Another question then becomes what role the work task situation and progression play in Taylor's Q4 compromise; the user might not perceive any substantial usefulness in applying Boolean operators because he or she is satisfied with results from queries where they are not used. Consequently, progress in a task is achieved without the power and control over searches mentioned by Hjørland (2015). Taylor's filter no. 2 'objective and motivation' and 5 'anticipated or acceptable answers' connects to work task situation, and relates to intentions in an IR situation as suggested by Ingwersen and Järvelin (2005, p. 204).

Finally, information need type as understood through Ingwersen's (2000) typology is seen as a critical factor in examining the user's process of formulating and modifying a query. The level of previous knowledge and connection to domain (for example, awareness of terms) combined with perception of work task and degree of variability of the need, is considered as providing the background for, and governing these processes.

2.5.1 Elements of a query

As a clarification, some definitions with regards to Boolean queries and their structures are given here. Using the building blocks terminology from Sormunen (2000, p. 33), queries are here described as composed of major facets (the concepts of interest), each represented through one or more terms (terms can be words or phrases). Terms that represent one major facet can be connected with the logical disjunction OR, resulting in facet queries (all terms concerning *one* of the major facets). The facet queries can then be combined using the logical conjunction AND, resulting in a complete query.

Also, this thesis is concerned with queries that incorporate Boolean operators as textual elements in a search string, submitted to an IR system. Built-in functionality such as using several search boxes in an interface, although often connected by default with Boolean AND, is not utilized in the data collection procedure.

2.6 Searching as a learning activity

The theoretical foundation presented in the previous section sees IR interaction as represented through cognitive structures that change over time. In a user-perspective, a connection could also be made to learning activities. Several reviews point to and discuss the relation between the act of searching for information, and learning (Marchionini, 2019; Rieh et al., 2016; Vakkari, 2016). In relation to this view, a possible framework in this thesis could

have been the Information Search Process (ISP) model developed by Kuhlthau (2004). Kuhlthau et al. (2008) reviewed research that used the ISP model to examine relevance judgements, and conclude that the stage of focus formulation in the ISP is of great importance in terms of information seekers' ability in "selecting pertinent information, developing more specific searches and becoming more critical of the found information." (Kuhlthau et al., 2008).

The data collection in this thesis is limited to one point in time, consequently it does not capture the participants' search processes in a longitudinal sense. Still, process stage within a search session is argued to be observable, and pre- and post-interviews could have provided insight into the participant's current stage within the ISP. The choice was however made to progress with the relevance interaction models presented in the subsequent section as frameworks of understanding, as the focus is on information needs and relevance criteria, and manifestations of different types of relevance.

2.7 Models of relevance interactions

In sum, several complex and debated concepts are to be included in a study that seeks to examine users' perception of the usefulness of certain search strategies, and how this perception might be linked to their information needs and relevance assessments. In section 2.1, Ingwersen's cognitive model of IR served as an illustration of the various dynamic components that are involved in IR interactions.

When focusing more closely on relevance manifestations during IR interactions, the stratified model developed by Saracevic (1996a) and extended by Saracevic (1997) is interesting, due to its depiction of levels or strata in both the system side and the user side of an information retrieval situation. At the time, the stratified model was proposed as an alternative to the system-oriented view, which focuses mainly on the system perspective and treats the interaction aspect of IR as implied, missing the details and nuances of such processes (Saracevic, 1996a, section 1) The stratified model consequently allows for a study of different manifestations of relevance and of how these can be thought of as interdependent across strata in the model (Cosijn & Ingwersen, 2000, p. 534). It can be considered as closely connected to the relevance manifestations proposed by Saracevic (1996b), and is even classified as a 'model of relevance' in a later publication by Saracevic (2017, p. 34).

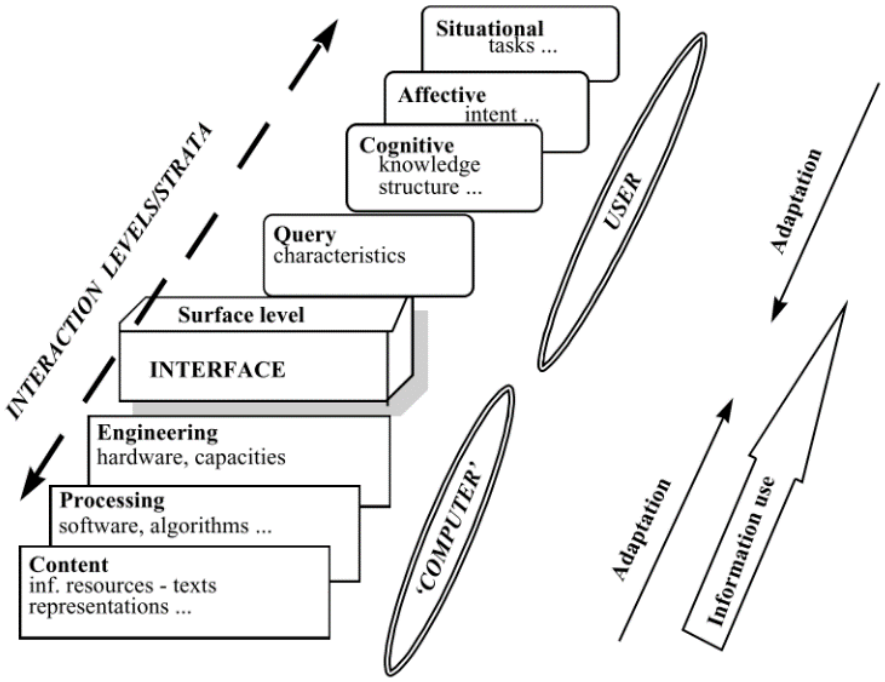


Figure 4: A stratified model of IR interaction. From Saracevic (1997)

The stratified model was later revised by Cosijn and Ingwersen (2000), who constructed a tabular overview of the relations between attributes and manifestations of relevance. A second version of this overview is described in Ingwersen (2000, p. 13). This second version is used as a framework to identify and understand relations between attributes and manifestations in the empirical data through analysis. The main purpose of this tabular view is that it suggests an understanding of how relevance manifestations assumed present in an IR situation can be characterized.

Attributes of Relevance	Manifestations of Relevance				
		⇔ Affective Relevance ⇔			
	Algorithmic	**Topical**	**Cognitive / Pertinence**	**Situational / Utility**	**Socio-Cognitive**
Relation	Query ⇒ Information objects (feature-based)	Subject/topic expressed in query ⇒ information objects	State of knowledge/cognitive information need ⇒ Information objects	Situation, work task or problem at hand as perceived ⇒ Information objects	Situation, task or problem at hand as perceived in socio-cultural context ⇒ Information objects
Intention	(a) System dependent (b) Intent/ motivation behind algorithm	(a) User /assessor expectations (b) Intent/ motivation behind query	Highly personal and subjective, related to information need, intentions and motivations	Highly personal and subjective or even emotional. Related to goals, intentions and motivations	Personal, subjective / org. strategy. Related to user's experience, traditions, scientific paradigms
Context	Tuning search engine performance (e.g. TREC)	All types of subjective relevance are, by definition, dependent (user´s/assessor´s			context-context)
Inference	Weighting and ranking functions	Interpretation of aboutness and subject matter at semantic level	Subjective and individualised process of cognitive/pragmatic interpretation, selection and filtering	User's ability to utilise information objects in a way meaningful to user	Users' (or group's) ability to utilise information objects, meaningful to environment
Interaction	Automatic relevance feedback or query modification	Relevance judgements are content dependent	Relevance judgements are content, feature, form & presentation dependent	Including interaction *with* environment	Including interaction *within* environment
		Increasing	**Time**	**Dependence**	⇒

Figure 5: Attributes and manifestations of relevance, from Ingwersen (2000, p. 13)

Borlund (2003, p. 915) illustrates the relationships resulting in manifestations of different relevance types in a visual model shown in Figure 5, adding another aspect to the understanding of relevance in IR situations involving users. For example, the 'cognitive space' component in the model is represented by both the participant and the author during the IR situations that generated the empirical data. This is due to the think-aloud approach chosen for the interviews, in which participant and author interacted in an asymmetrical, conversational manner, with the author asking the participant for reflections and

explanations. The motivation for the choice of interview technique is elaborated in section 4.5.

Query versions are also depicted in Borlund's model, which fits well with the research design involving two versions of queries for each participant. The resulting complexity involving different manifestations of relevance is shown here as numerous, concurrent relationships present in an IR situation. The empirical data and subsequent analysis in this thesis aim at providing a glimpse of this complexity.

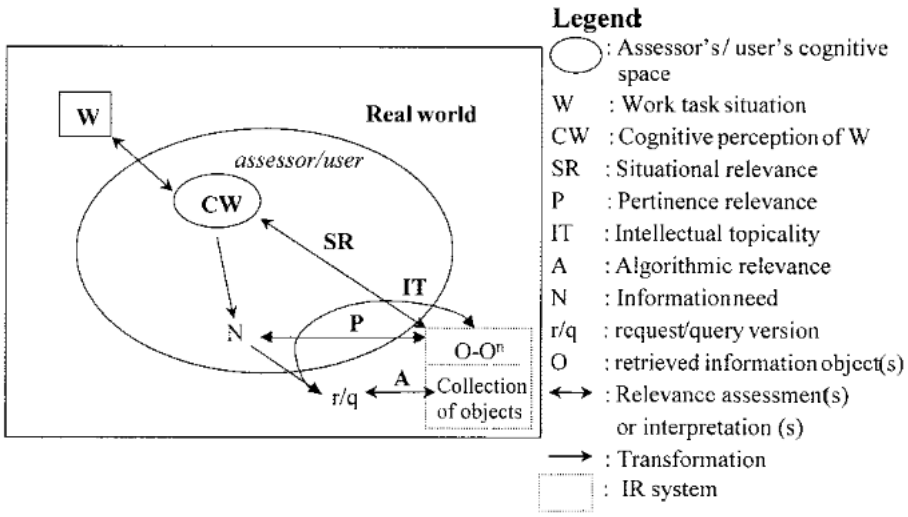


Figure 6: Main types of relevance relationships involved in an IR situation, from Borlund (2003, p. 915)

Returning to the research questions with the preceding theoretical foundation in mind: RQ1 is concerned with *how* the user group in question perceive the usefulness of Boolean queries. The Cognitive viewpoint presented in section 2.1 is chosen as the basis for understanding the *how*, because it provides a research framework in which the participants' information need plays a central role. This individual and dynamic information need is considered an essential context if one is to examine a user's perception of usefulness of a given retrieval technique.

RQ2 seeks to identify the relations between this perception of usefulness, and the participants' information needs and relevance assessments. The background for RQ2 is an assumption that a compromise must be reached between information needs and queries, as

indicated in the research literature (Byström & Kumpulainen, 2020; Taylor, 1968), and that information need types and label effects (Ingwersen, 2000) might influence the process in which this compromise is reached. A second assumption is that different manifestations of relevance can be identified through the observation of relevance criteria stated by users, by examining the relationships suggested by Ingwersen (2000) and Borlund (2003).

The remaining text is organized as follows: In section 3, previous research considered relevant to this study is reviewed, before the research design and methodology is presented in section 4. Section 5 presents the results, with a discussion following in section 6. In section 7 a conclusion is drawn, and future research efforts are suggested.

3 Previous research

Researching students' experiences with and perceptions of Boolean queries involves several wider research themes within information science, among them information seeking behavior, information needs, relevance assessments and query quality. With regards to information seeking behavior research, a pragmatic choice is made in looking at a selection of studies that examine the search context and search habits of students in higher education. Information needs and relevance assessments are treated more at length, as empirical research on these themes informed the chosen research design to a great extent.

3.1 Higher educational students' search context and search habits

In a meta-synthesis of research on graduate (doctoral and master) students' information seeking behavior, Catalano (2013, p. 267) found that few students within this group utilized advanced search techniques such as Boolean operators. In addition, one of the general conclusions of the synthesis was that empirical research projects should distinguish between master and doctoral students as groups during data collection, due to differences in search habits and behavior (p. 269).

Dempsey and Valenti (2016) analyzed search logs from a discovery service system and found that few of the 118 college student participants in their study (15%) used Boolean operators to edit their queries. In addition, participants seemed to struggle with the combination of different operators. Although these findings are based on a small population, they provide some indication as to usage of operators in discovery system settings.

Cmor and Li's (2012) case study investigated what impact discovery services in academic libraries might have on pedagogical approaches to teaching information literacy. They argue that the introduction of such services makes it possible to save time in library instruction classes previously spent on teaching students Boolean search strategies, search rules and interface related details within separate databases. They further contend that due to the simplicity and familiarity of searching a discovery service, time can rather be spent teaching students how to engage with information and critically assess sources, thus increasing the focus on students' critical thinking skills and making a change to an exploratory pedagogical approach rather than explanatory pedagogical approach (2012, p. 451). In the years that have passed since Cmor and Li's study, discovery systems have been steadily implemented in a growing number of institutions globally, becoming "the de facto search tool in many libraries" (Hamlett & Georgas, 2019, p. 231).

In a Norwegian higher educational setting, Østbye's master thesis (Østbye, 2018) provides insight on students' information needs when searching the discovery service Oria. The findings showed that participants who reported using advanced search techniques such as Boolean operators, were graduate students or had attended library instruction classes (Østbye, 2018, p. 75), and that when Boolean search was used, this was done with few terms.

3.2 Quality of queries and query-terms

In a longitudinal case study, Pennanen and Vakkari (2003) examined relations between students' existing knowledge on their chosen research topic and their search experience, on search tactics and outcomes. Boolean searches were done by the students in the subject database PsycInfo. One of the findings in this study was that an increasing number of search terms used, had a positive effect on the number of partially relevant results when the students were further along in the task (Vakkari et al., 2003, p. 452). From this, it can be interpreted that increased subject knowledge combined with increased number of query terms can indicate that more partially relevant results are identified. In a related study, Pennanen and Vakkari (2003) found that the more familiar students were with their chosen topic, the more capable they are were of distinguishing useful terms when assessing results (p. 765).

Lee and Chung (2016) explored students' relevance assessments and employed basic and advanced Boolean searches in the study design. Their goal was to compare the effectiveness of discovery services as opposed to subject databases. This comparison of different kinds of IR systems is outside the scope of this thesis project, but Lee and Chung's study is interesting due to the choice of research design and methodology. Basic and advanced version of queries for each were constructed and run in the databases and discovery tools that were to be compared, and results were assessed by graduate students who assigned graded relevance on a 4-point scale, whereby modified recall and precision measures were calculated. This was consequently done in an intellectual topicality manner, as the queries were not based on the participants' needs.

The previously mentioned study by Lowe et al. (2018) examined the potential benefits (in terms of topical relevance of results) of teaching Boolean search techniques for first-year undergraduate students, as opposed to using natural language searching. Boolean and natural language versions of queries for three sample topics were carried out by the researchers across eight databases (including Google Scholar, deemed a database for the purposes of the study). Relevance assessments were then conducted using the rubric shown in Figure 7, page 33. The basis for this rubric approach is that the top 25 results returned by the queries were assessed based on title and abstract. The six co-authors formed three pairs, who then used the description of each category in the rubric to determine a relevance score in terms of presence of concepts in the retrieved results. The degree of presence of concepts corresponded to the categories 0 – not relevant, 1 less relevant, 2 relevant, and 3 very relevant.

The resulting analysis included the number of results that were placed in each of the four categories, average score for each result set, and the degree of overlap between results of both queries in each database. Their study "found no clear advantage in relevance of results between natural language and Boolean searching[…]" (Lowe et al., 2018, p. 531). The authors conclude that these findings suggest that for introductory courses, librarians can spend more time "on other, more substantial information literacy concepts such as topic and question development (including search terms and terminology) and source evaluation" (Lowe et al., 2018, p. 531).

TABLE 1 Article Relevance Evaluation Rubric			
Not Relevant (0)	**Less Relevant (1)**	**Relevant (2)**	**Very Relevant (3)**
0 of total concepts represented OR false hits, terms are there but used in different ways (e.g., social work instead of social rejection)	Less than half concepts represented OR concepts are there but not relevant to research question	Majority or all of concepts represented either in title or abstract but when looking at abstract, may be tangential to research question	All concepts represented in title or abstract and abstract is relevant

Figure 7: Rubric approach to assessing relevance (Lowe et al., 2018, p. 521)

A related study by Lowe et al. (2020) sought to investigate whether simple Boolean search with the operator AND is sufficient for advanced upper-level undergraduate or graduate students doing research for a master thesis project or similar, or if relevant literature would be missed by only using this operator. Using descriptions deemed to represent five realistic search scenarios for the user group in question, simple (using the operator AND) and advanced (using the operators AND and OR, as well as truncation) queries were constructed by the authors. However "The overall goal was not to create perfectly crafted searches but to try to reflect search behaviors that would be typical for undergraduate researchers." (Lowe et al., 2020, p. 3).

Eleven databases (Google scholar included) were searched. The first 25 results in each database were assessed for relevance by analyzing title and abstract, using the 4-point rubric approach introduced in Lowe et al. (2018). Their results showed that "Based on relevance, there is no compelling evidence that either search is superior" (Lowe et al., 2020, p. 5), although it is stated that the low level of overlap within certain databases warrants that performing both simple and advanced Boolean search is advisable for students that seek comprehensiveness.

A similar approach to the rubric approach used by Lowe et al. (2018; 2020), called 'aboutness measure' (shown in Figure 7) is utilized by Toms et al. (2005), in an effort to identify manifestations of topical relevance as defined by Saracevic (1996b).

Table 2. Aboutness Measure

Code	Definition
5	pages directly related to the topic and containing clear info on the topic,
4	pages that provide some information that is related, or leads directly to the answer
3	pages that about the topic but may be broader or narrower that the topic
2	tangentially related but not really in the topic area
1	pages that are clearly not about the topic at all

Figure 8: Aboutness measure used by Toms et al. (2005, p. 69).

In Toms et al.'s study, external assessors used the measures to assess the topical relevance (as to the used query) of web pages that were judged as relevant by participants (Toms et al., 2005, p. 69).

A discussion of the rubric approach used by Lowe et al. (2018; 2020) is appropriate here, as it was one of the issues that made the author interested in looking into similar research questions. The operationalization of relevance in Lowe et al. (2018; 2020) corresponds to the type of 'topical relevance' defined by Saracevic (1996b, section 4) as the "relation between the subject or topic expressed in a query, and topic or subject covered by retrieved texts [...]". The employment of a graded scale represented by the rubric approach highlights some of the interesting questions that arise when relevance is assessed and defined in studies that do not involve 'real' users as participants and assessors. In the rubric approach, relevance is assessed according to a static topical request, not a dynamic information need belonging to a user. Consequently, the assessments do not consider the situational context of such a user, e.g., his or her existing knowledge of the research theme in question or intended use of the information found.

The perspectives on Boolean searching represented in the reviewed literature so far highlights the challenge of balancing higher educational students' expectations of relatively easy and effective searching through discovery tools, and the comprehensiveness in, and potential control over search results represented by Boolean queries. In ascertaining the quality of different levels of complexity (simple/advanced) in Boolean queries, topical relevance of results has been used as a performance measure in empirical studies. How

relevance is conceptualized and used as such a measure is dependent on research tradition and associated theoretical viewpoints, as discussed in section 2.3.

3.3 Studies of relevance assessments and criteria

The methodological and analytical approach in this thesis is inspired by a handful of empirical studies on relevance assessments and criteria, which will be reviewed next.

Spink et al. (1998) examined what factors that affect users' graded relevance assessments, focusing on the role of partially relevant results. Findings from four separate studies where users perform their initial searches on an information need are reported. One of the conclusions were that partially relevant results can be of importance to searcher who are at an early stage in a task, in that the results enable a development in terms of understanding the topic in question (Spink et al., 1998, p. 612).

3.3.1 Relevance criteria beyond topicality

In a study by Park (1993), participants were asked to described what criteria they relied on when assessing the relevance of results from a search based on their own information needs. Interview transcripts were analyzed for recurring themes, resulting in what Park summarized as citation-based and user-based characteristics, which then both contribute to 'user-based-relevance' (Park, 1993, p. 342). Citation-based characteristics concern the elements of the retrieved information objects, such as author, title or abstract. The user-based characteristics are described as belonging to three interconnected groups: 'internal context' (i.e., the user's previous knowledge), 'external context' (i.e., search quality or search goal) and 'problem context' (the user's present problem and how information affects its construction and change) (Park, 1993, pp. 341-342).

Park's (1993) categorization of characteristics involved in relevance assessments is strongly connected to the cognitive viewpoint of IR, as the user's situation and context is emphasized. Further, links can be argued to be seen between the three contexts described by Park, and the relevance manifestations proposed by Saracevic (1996b): internal context may be seen as related to cognitive relevance, external context to motivational relevance, and problem context to situational relevance.

Barry (1994) examined what relevance criteria users relied upon beyond that of topicality. In her exploratory study, the participants' individual information needs provided the basis on

which references were retrieved and assessed. The assessments were then analyzed and coded using content analysis, resulting in a list of identified criteria which were then grouped. The resulting overview of mentioned criteria by participants highlighted the situational aspects of such assessments, both in terms of the participants' background and experience, and in terms of information needs.

An important difference between the methodology in Barry's study and the present thesis is that in the former, relevance assessments were made outside of context of the IR system. In addition, the participants were shown a random sample of 15 document representations gathered from the search, of which three were randomly chosen and presented in full text as well. While providing a stable number of assessed documents per participant, the data does not directly reflect how participants would select and assess retrieved results in an interactive IR setting. This interactivity, which includes the reformulation of queries and participants' own choices of results to assess, is seen as crucial in providing answers to the research questions in this thesis.

3.3.2 Influence of task stage and topic knowledge on relevance criteria

Vakkari and Hakala (2000) studied the relation between changes in relevance criteria and changes in task process, thereby incorporating the role of the user's accumulated subject knowledge as a result of the search process. In their exploratory study, 11 graduate students were asked to do searches at three different stages during a research project and assess the relevance of retrieved results at each stage. The categories of relevance criteria identified by Barry (1994) were used as a starting point to distinguish the different criteria stated by users. Graded relevance assessments were also included in the data collection, which allowed for an analysis of the distribution of such assessments over time. Vakkari and Hakala found a connection between task process stages, and the user's capacity to construct queries with precision:

> At the beginning of the task performance, the conceptual construct representing the task is undifferentiated and the relation between the central categories is fuzzy. The user is unable to formulate exact queries because he or she is unaware of concepts and relations. (Vakkari & Hakala, 2000, p. 559).

In other words, the effort needed to reach a compromised Q4 need can be viewed as large in these cases. The authors emphasize the role of the task in relevance assessments: “the relevance of documents is assessed by the actors in terms of their support and contribution to a certain task” (Vakkari & Hakala, 2000, p. 541). An assumption could be that by extension, a query technique is perceived as useful if it generates results that helps the user in accordance with an ongoing task. This can be used as an argument for focusing on situational relevance of results in understanding the degree of perceived usefulness of Boolean queries.

In a comprehensive analysis of a previously gathered dataset, Taylor et al. (2007) examined the relation between stage in a search process and what relevance categories participant relied on. This was done via a coding process through content analysis of the data. The most frequently stated criteria among the participants was that of ‘Specificity’, defined as “Expressions of specificity: specific/general; technical; too technical; not technical” (Taylor et al., 2007, p. 1077). It is interpreted here as to what degree the topical content is specified. ‘Specificity’ was followed by ‘General topicality’ as the second most frequently stated. The findings suggest that criteria pertaining to topic play an important role.

Taylor (2012) explored how users’ choice of relevance criteria were related to stages in Khulthau’s ISP model, and whether users preferred certain criteria depending on stage. This study analyzed frequency of relevance criteria within quantitative data. Data were derived by asking the participants to self-report via predetermined categories, what distinct criteria they relied on during assessments, as well as stage in the search process while interacting with an IR system. Both criteria categories and stages in ISP categories were presented to participants through a customized version of a commercial search engine, allowing them to report non-verbally on these items for each result while searching on their own. Participants could submit data multiple times through the web page. The data collection dealt strictly with results that were judged as relevant by participants, leaving out criteria for partial or non-relevance. While collecting data unobtrusively from participants interacting in an IR system setting, data collection relied entirely on participants’ interpretation of, and their choices between, a relatively large number of alternatives (15-20) for each result.

Using a similar methodology, Taylor (2013) examined the influence of work task on choice of relevance criteria. Work tasks were categorized according to assignments given to

participating students as part of a class. The findings indicated strong statistical relationships between what criteria used to assess relevant and partially relevant results, and work task (Taylor, 2013, p. 536). Similar to Taylor (2012, 2013), this thesis aims to connect the tabular overview of interactions between attributes and manifestations of relevance proposed by Cosijn and Ingwersen (2000) to user-centered relevance criteria studies such as Barry (1994). This connection inspired the analytical approach taken in this thesis.

3.3.3 Relevance manifestations

As mentioned in the introduction section, the application of non-user, topicality-based assessments by Lowe et al. (2018; 2020) inspired the author to examine how a user-centered study on the subject of Boolean queries could be conducted, and how other manifestations of relevance might be involved when looking into similar research questions.

In a study reported by Cosijn (2006), criteria for relevance were derived via a survey, from users that were in the final stages of completing a work task. Predefined, research-based categories of criteria are mapped to the relevance types described by Saracevic (1996b), providing the basis for an analysis of relations between criteria and work task, and criteria and relevance manifestations. This mapping procedure is interesting, as it directly operationalizes criteria to relevance type. However, this seems to add less emphasis on possible dynamics with regards to criteria that are used concurrently, or that rely on each other.

In a synthesis of empirical research on user relevance criteria, Saracevic (2017, pp. 57-58) identifies clues and criteria discovered in empirical research as, in a generalized manner, belonging to the following groups: Object characteristics (content, object, validity) and human characteristics (use or situational match, cognitive match, affective match, belief match). An interdependency between the first three and the last four groups of relevance criteria during assessments is emphasized (Saracevic, 2017, p. 58). An earlier parallel to this grouping is seen in the object- and user-characteristics found by Park (1993). Saracevic's grouping of categories (2017, pp. 57-58) provided the background for the deductive coding procedure described in detail in section 4.10.1.

4 Methodology

This section describes two processes: One concerns the decision on a proper set of data collection instruments and experimental components with which to examine the research questions. The other concerns decisions regarding the research questions themselves. As will be shown, pilot testing of a preliminary design revealed that certain changes had to be made to the research questions that originally were intended to be examined at the outset of this project. These changes will be accounted for in the subsequent sections.

4.1.1 Research design

The table below, based on Mason (2018, p. 26) illustrates what data sources and methods were considered for the preliminary qualitative design to be applied in the pilot test, and how these were intended to help answer the initial research questions:

Research questions	**Data sources and methods**	**What could they yield?**
1. How do graduate students within the educational sciences perceive the usefulness of advanced Boolean queries, when searching a subject database?	Graduate students: Participation in mediated searches Concurrent or retrospective think-aloud interviews Screen recording of two search sessions	Statements about perceived usefulness of searches utilizing Boolean operators during an IR situation. Statements regarding information need before and during searching. Relevance assessments of results from one search session with user-generated queries and one session using Boolean operators to expand the last version of the initial query Number of search terms used in the queries Number of total results for each search.

2. How is the perceived usefulness of advanced Boolean queries among students related to the students' information needs and relevance assessments?	Interview statements: Thematic analysis of interview transcriptions Interview statements: Analysis of frequency of non-relevance/partial relevance/high relevance type of assessments and relevance criteria within and across sessions.	Descriptions of students' information needs. Frequency of assessments pertaining to non-relevance/partial relevance/high relevance and relevance criteria Degree of correspondence between relevance assessments and stated perceptions on the usefulness of advanced Boolean searches within sessions and across participants.

Table 1: Initial research questions and proposed data sources and methods

The initial research questions shown in Table 1 suggested that a qualitative, user-oriented methodology would be appropriate. However, attitudes and experiences regarding the use of Boolean queries among students might also have been examined in a more quantitative manner using a survey, perhaps combined with open-ended questions. Data from such a survey could for example offer insight into how many of the respondents who regularly used Boolean operators when constructing queries, and what potential barriers they experienced constructing such queries. The level of usefulness the students experience with using Boolean operators might also have been examined, through a graded scale rating or a similar measure.

The survey approach was considered at length, especially after a long period of poor recruitment in the project. Still, the ambition was to involve users engaged in the moment of an actual information searching process, with a focus on the role of relevance and characteristics of information needs in such situations. The choice was therefore made to carry on with a qualitative research design.

As can be seen in Table 1, the preliminary research questions concerned users' perceived usefulness of advanced Boolean queries that expanded an initial, user-generated query. The pilot test research design therefore specified that the participants' information needs were

to be represented by their own queries in the beginning of the session. Participants were to be kindly asked to search the ERIC database via the EBSCO*host* interface, based on what kind of information they were looking for at that moment. Following this interaction with the database, and possible refinement by the user because of relevance feedback, the final user-generated queries were then to be developed into more elaborate Boolean queries by the author and participant in collaboration and tested and reviewed by the participant. Relevance assessments and stated criteria for the 10 first results were then anticipated to implicitly provide insight into the perceived usefulness of the queries**.** The initial design was chosen to enable exploration of the perceived usefulness of an expanded version of a query through using Boolean operators, as viewed by the students. A limit of assessing the 10 first results in each query was intended to reduce the burden on participants.

The instance of a single search session based on one stated information need per participant is understood as an example of 'session-based interaction' with an IR system, in which one might "expect to observe the dynamism and variability of the perceived information need and search task owing to interactive processes of learning and cognition[...]", while the user's perception of work task can be viewed as quite stable (Ingwersen & Järvelin, 2005, p. 303). The influence of these interactive, cognitive processes on perceived usefulness of query techniques is of particular interest. While the participants' work task (completing a master thesis) could be assumed to be stable throughout the session, each participant's information need might develop and change in nature during the duration of the session, due to the encountered information objects, through searching and learning. The relevance assessments of the results in a second query in a session setting, will quite naturally be affected by the results viewed in the first (Ingwersen & Järvelin, 2005, p. 353).

As pointed out by Sormunen (2000, pp. 31-32), the interface/intermediary component in the cognitive IR model (Figure 2, p. 10) is considered as a cognitive structure engaged in query formulation. In the research design, these structures are present through the author acting as the intermediary, and in the retrieval functions underlying the EBSCO*host* interface. This point is elaborated in section 4.7.1.

The view in this thesis is that information needs are situational in nature, and that they are inseparable from an ongoing task. The is to provide context to the interpreted need, although a humble view is taken in terms of the ability of the author to truly create queries

that can relate well enough to participants' information needs, both in the initial and the final research design. A compromise is always present, and the information needs and queries are not considered as equalized.

Thorough pilot testing is recommended in the IIR methodology literature (Kelly, 2009, p. 60), and was considered necessary in order to gain experience with collecting data, and detect possible problems with the chosen methods. Data for the pilot test was gathered between February and March 2021. The pilot test is reported in detail in section 4.6.

4.2 Selection and recruitment

Participants were recruited among graduate students enrolled in an educational science program at a university in Norway. Students in their final part of a 5-year graduate degree were selected as participants. At this point in their degree, they had been instructed by faculty to start searching for literature as part of preparing for their master thesis. This was deemed by the author to increase the chances of recruiting students who were actively searching for information for these purposes. In this way they represent a purposive sample, as defined by Kelly (2009, p. 67). The students had over the course of their studies been introduced in library instruction sessions to use both the university library discovery service, an Exlibris Primo installation called Oria, and academic subject databases when searching for literature. It is important to note that all participants were at an early point in their master thesis project at the time of recruitment.

A total of 241 students met the criteria and were designated to be in the target group for recruitment. An invitation to participate was posted at the students' digital learning platform, in which the university library had established a dedicated space with digital resources. Recruitment was also carried out in cooperation with faculty staff, via email invitations that were sent to and forwarded by faculty who were responsible for courses that ran parallel to the master thesis course the same semester. In the email invitation, a brief description of the project was given. A separate, more detailed information sheet was included as an attachment. The invitation contained a link to an online form, in which the students who wished to participate were asked to provide the following:

- Name and email address
- Consent to participate in the project, confirming that they had read the information sheet
- A brief written description of what kind of research they are searching for at the present time, in relation to their master thesis
- A query using terms of their own choice, as they would have planned to use, when searching the subject database ERIC
- Possible dates and times for participation in the study, chosen from a predefined list

See Appendix A for the online form. Upon receipt of the online form, the author contacted each participant via email to schedule an appointment for data collection. The collection of personal data in this project was approved by NSD January 26th, 2021. See Appendix H for the NSD approvement.

The students were also invited to participate in the study, by the author in person during breaks in separate online classes hosted by faculty and when attending an online workshop on literature searching hosted by the university library. The workshop consisted of a 2 x 45-minute introduction on how to do a literature search for the master thesis, followed by online group conversations between students, with the option of getting guidance on literature searching by a university librarian. The presentations included information on how the Boolean operators AND, OR and NOT function, and how keywords and synonyms for a research topic can be developed and combined into queries. It concluded with an overview and repetition of previously introduced features of the university library's discovery service, Oria. The database used in the data collection for this thesis, ERIC via EBSCO*host* was not presented during the workshop, but students could access information about how to search this database on the digital learning platform.

4.3 Ethical considerations

The author participated as one of the speakers during the presentations in the workshop and was available for guidance during the group conversations. The author is however not associated with the educational science study programs on a day-to-day basis. There is of course still a possibility that participants might have been adapting their responses during data collection, because they knew the author is associated with the university library. In the

previously mentioned information sheet, it was clearly stated that the research project was affiliated with Oslo Metropolitan University, and that participating was not a part of the online workshop of their own university.

It was also underlined that choosing to withdraw from the study at any time had no negative consequences for the students. To ensure that all participants had similar information at the time of participation, final data collection was carried out after the workshop took place. The following information about the participants is not included, to ensure anonymity: University affiliation, exact graduate degree program, gender, and age.

4.4 Experimental setting

To accommodate for a natural but controlled environment for the data collection, a temporary lab on campus was to be established. However, due to the outbreak level of the Covid-19 pandemic at the time, the decision was made to collect data digitally. Although participants had access to campus facilities, all classes and workshops were kept digital as part of an effort to reduce the outbreak during the months in question. Data collection in a physical lab would involve different participants using the same computer and visiting campus more than necessary and was therefore considered to be inappropriate, and last but not least in breach with sound research ethics, given the outbreak situation. The resulting digital experimental setting described in the following was used for all instances of data collection in this thesis.

Data collection sessions were scheduled with each participant using the video meeting service Zoom[4], licensed via the author's Oslo Metropolitan University affiliation. An email with a link to the Zoom meeting was sent each participant based on their preferred date and time. Each Zoom-meeting used a generated unique meeting ID to prevent anyone other than the author and the invited participant to gain access to a single session. Cameras in Zoom were activated for both the author and participants during the entire data collection session. No video from the cameras were recorded. During the interviews, sound was recorded using the NSD-approved 'Nettskjema diktafon' app. This was installed on an iPhone and an iPad, using the iPad as a back up recording device. The use of Nettskjema diktafon recordings via devices outside of Zoom ensured encrypted storage of interview sound data.

[4] https://zoom.us/

Screen recordings were done to provide data on the on-screen interactions. To prevent personal data from being collected in screen recordings, control over a shared computer screen was given to participants using the 'remote-control' function in Zoom. This enabled participants to interact with the database using their own computer, while the screen on the authors computer was being recorded, using Screencast-o-Matic[5]. The recordings also served as support for remembering details when transcribing and analyzing the interview data later on. Figure 9 shows the digital data collection set up:

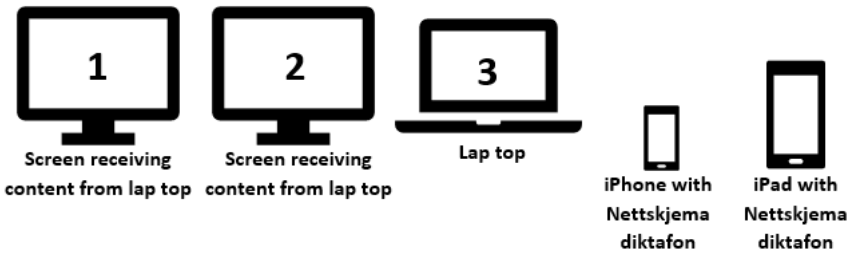


Figure 9: Digital data collection set up

The Zoom application with video and audio ran on the laptop with webcam [3], while the shared, participant-controlled screen showing the ERIC database ran on [1]. The content of screen [1] was recorded using Screencast-O-Matic. Screen [2] contained the authors working documents during data collection, such as the interview guide and documents with information about queries.

The range of possibilities and limitations of this fully digital experimental setting were unknown at the outset of the data collection. Extensive pilot testing revealed that strict procedures had to be established to get all technical components to work right and at the right moment during search sessions. Examples of this were the placement of recording devices with regards to distance to speakers, step-by-step routines for starting and ending both the sound and screen recordings, and the logistics of handling shared screen content and giving control in Zoom to the participant. These actions had to be practiced in advance of search sessions so that they required a minimum of attention from the author, ensuring a

[5] https://screencast-o-matic.com/

focused and well-functioning interview situation. Procedures are documented in the protocol shown in Appendix B for the pilot study, and in Appendix D for the final data collection.

4.4.1 ERIC via EBSCO*host*

Several alternatives were considered when it came to choosing an IR system in which carry out the searches. Even though the Exlibris Primo installation Oria was assumed to be well known to the user group in question, the research questions and resulting design limited the scope to subject databases. One major reason for this choice was the possibility of examining participants' satisfaction with results in a subject specific collection that they perhaps did not use as often as Oria, but still are encouraged by faculty to use. Another was the availability and detail of system documentation of such databases, which allowed for broad insight into search functionality.

Given that the participants were graduate students within the educational sciences, searches were chosen to be carried out in the ERIC database, consisting of bibliographic references with and without full text content within the field of education research. The database includes content from peer reviewed journals and non-journal content such as books, proceedings and grey literature from 1907 to the present (Institute of Education Sciences, 2021). For the user group from which participants were recruited, the ERIC database is searchable via two interfaces. In its original interface available at https://eric.ed.gov/, and via EBSCO's EBSCO*host* Research Platform, through which several of the university's subscribed databases are available.

The fact that multiple databases are available for the students through the EBSCO*host* interface was considered as an argument for choosing this interface. Gusenbauer and Haddaway (2019)'s exhaustive review of retrieval qualities among IR systems shows that the EBSCO*host* version of ERIC is superior to the original version in terms of search functionality, especially with regards to functional exact phrases in queries, and available advanced search field. The way each submitted query is presented in this interface, and how this presentation is combined with number of search results for each search was also a strong argument for this choice. This provided the author and the participants with easy access to information about each query during the think-aloud interviews.

Below, some of the key aspects of the settings and search functionality in ERIC via EBSCO*host* are presented. These are all considered to be essential parts of the experimental setting in that they affect how the queries operate, how algorithmic relevance is calculated through i.e., weighting of different criteria, the number of displayed results, and what metadata are displayed for each result. This in turn does have effects on relevance assessments and perceived usefulness of queries among participants. While this thesis does consider participants' IR system interaction, the focus is on perceived usefulness of queries through satisfaction with results. Hence it is not a usability study as such, but an IR study.

Search screen

Both the basic and advanced search screens in EBSCO*host* offer Boolean functions: "The Basic and Advanced Search Screens let you create a Boolean search with operators such as AND, OR and NOT. These searches are not designed to yield significant results if you enter long phrases or questions in everyday English" (EBSCO Connect, 2018). The advanced search screen is the default start screen when the EBSCO*host* interface is accessed through the university library web page. All searches were carried out through the advanced screen in both the pilot tests and the final data collection. Figure 10 is a screenshot showing retrieved results in the EBSCO*host* ERIC interface in the advanced search screen. The 'Peer reviewed' limiter has been activated, generating a second set of results displayed as 'Search ID# S2'.

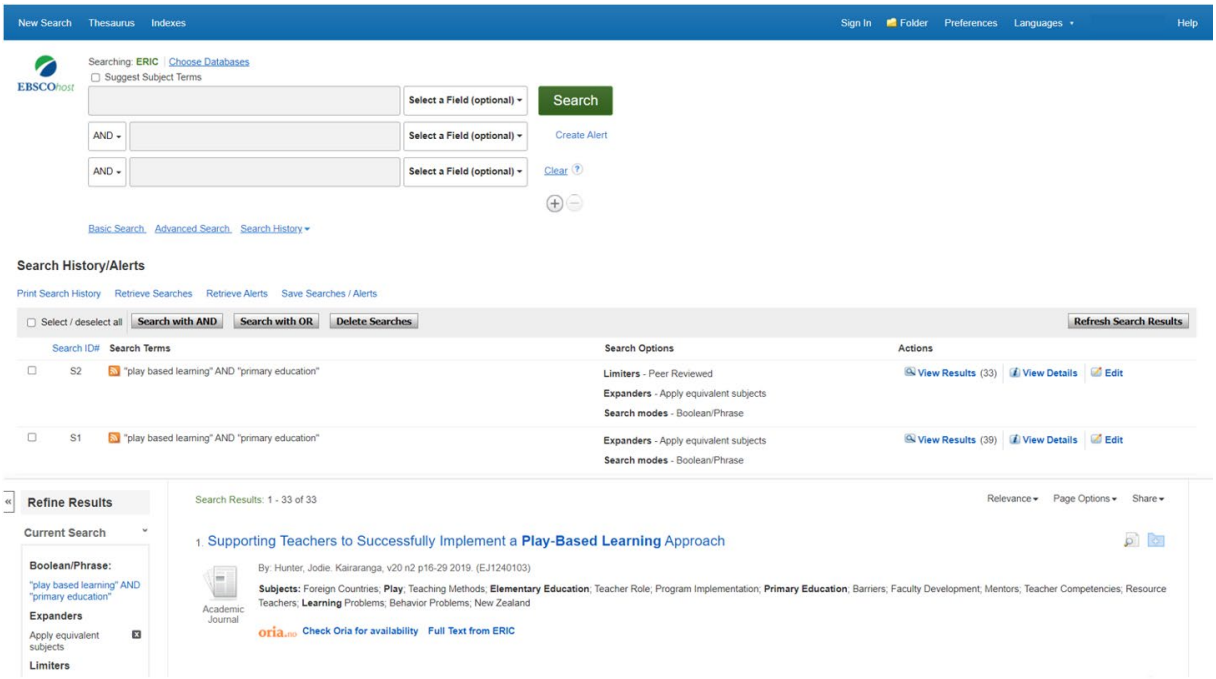


Figure 10: Retrieved results in the EBSCOhost ERIC interface.

Search fields and search mode

The search mode 'Boolean/Phrase' was used for all queries. This mode is defined with the following functionality: "Supports any Boolean searching or exact phrase searching. Stop words are ignored when part of phrases being searched" (EBSCO Connect, 2021). Search field was not specified during searches. This meant that the terms used in the queries were searched in these default fields: "Title, Author, Subjects (Descriptors/Identifiers), Institution Name, Core Subjects, and the Abstract Summary." (EBSCOhost, n.d.-a).

Expanders

The expander function named 'Apply equivalent subjects' was left as activated, the reason being that it is so by default when students access the database from the university library web page. In doing so, the performance of the Boolean queries reflects that of the real-world setting for the user group in question. This expander function "[...]uses mapped vocabulary terms to add precision to unqualified keyword searches. (An unqualified keyword search is a general search that does not specify a field to search, such as title, subject or abstract.)" (EBSCO Connect, 2019). As search field was not specified, the queries were treated as 'unqualified keyword searches', potentially activating the expander. There is no

information displayed in the interface that indicates if the expander has been activated or not, when a set of results is retrieved.

Ranking of search results

In EBSCO*host*, result lists are by default ranked according to relevance. EBSCO gives some insight on which criteria are used in this ranking: "The major contributing factor in relevance scoring is the frequency of the user's search terms in matching EBSCO*host* metadata and full-text records. Like all search engines, EBSCO*host* begins by finding records that contain the words that match the user's search query" (EBSCO Connect, 2020). The fields subject heading, title and author-supplied keywords are stated as the most influential in this process. A full disclosure of functionality is not provided.

An expander function is also active in the relevance ranking of results:

> Enhanced Subject Precision utilizes mapped vocabulary terms from multiple sources and users [*sic*] natural language to add precision for topical searches. When a user's search term matches a known concept, records about the concept receive an additional relevance boost. Multilingual search queries are also mapped to increase opportunities for concept matching. (EBSCO Connect, 2020)

It is not stated whether or not this is the same functionality as described for the 'Apply equivalent subjects' expander. All filtering options of results, such as limiting to certain publication years, or only viewing peer-reviewed content and similar functions were available to participants as usual. Number of displayed results per page was left at the default 50 in all sessions.

4.5 Approaches to interviewing

Both semi-structured pre- and post-search interviews and think-aloud interviews were utilized during data collection. A description of each of these elements is provided here.

4.5.1 Think-aloud approaches

Both concurrent and retrospective think aloud-approaches to interviewing were considered for the parts of the data collection procedure where the aim was to elicit participant's relevance assessments and other interaction-related thoughts that occur during the IR task performance. The concurrent approach requires that the test participant states his or her thoughts and reflections to the researcher while performing the given tasks in an experiment, thinking out loud. The recording of real-time responses from participants can be seen as a advantage (Sanghee Oh & Wildemuth, 2017, p. 201) but this interview form can also be quite demanding and even awkward for participants in IIR experiments, often because the tasks are new, or the search interfaces that are used might be unfamiliar (Kelly, 2009, pp. 84-85).

The retrospective approach on the other hand, lets the participant perform the IR tasks on his or her own, unobserved. It then requires the participant to review his or her interactions with the IR system(s) in question, in dialogue with the researcher. Participants are asked about thoughts and reflections regarding their own search behavior and decisions while watching a screen recording of their on-screen actions, or a transaction log (Sanghee Oh & Wildemuth, 2017, p. 201). The retrospective approach might yield better data because the participant's focus is undivided, as opposed to in the concurrent approach (Kelly, 2009, p. 85).

Finding a practical way of combining a retrospective approach with the digital experimental setting proved quite challenging. On one hand, the participant could be 'left alone' in the digital test room by the author leaving the actual physical room, and/or turning off the camera in Zoom. An agreement could be made beforehand on a set time, or the participant could reach out when he or she felt finished with the search session. This was considered quite a cumbersome arrangement, and one that would result in an unnatural setting for the participant, who in any case would probably not feel 100 percent on his or her own as long as the Zoom meeting still was ongoing. It would also make it difficult for him or her to reach out in case of questions, technical problems, or any other issues. Another downside to the retrospective approach is that it would increase the total time spent significantly, putting a large cognitive load on the participant.

It was the author's belief that searching and assessing the relevance of results in what was expected to be a known search interface to the participants, represented a somewhat familiar IR task. While the concurrent think-aloud approach was expected to be cognitively quite demanding for participants, the limited degree of interaction across multiple websites or systems combined with shorter time spent in sessions spoke for this alternative.

The concurrent approach itself can be performed in different ways with regards to methodology. What is now described as classic think-aloud (CTA), originally developed by Ericsson and Simon (1993), implies that strict procedures are to be followed to ensure validity of the method (McDonald et al., 2015, p. 388), especially when it comes to what kind of interventions a researcher might make during data collection (Makri et al., 2011, p. 342). A modification of the CTA permits such interventions by the researcher, resulting in a 'relaxed' or 'interactive' think aloud (ITA), in which the interventions are intended to increase the chances of eliciting explanations and experiences from participants, rather than more procedural descriptions (McDonald et al., 2015, p. 388). The value of ITA over CTA in getting verbalizations that concern explanations and assessments is also emphasized by Hertzum et al. (2015).

The goal of eliciting explanations in terms of which criteria participants relied on spoke for the use of ITA rather than a strict CTA. While the ITA approach described by McDonald et al. (2015) is originally examined in the context of usability studies, it is considered to appropriate to the research questions examined in this thesis.

4.5.2 Pre- and post-search interviews

In both the pilot test and the final data collection procedure, a pre-search interview was carried out in addition to collecting think-aloud statements. It utilized a semi-structured guide that allowed for follow-up questions. The main purpose of the pre-search interview was to get a clear understanding of the information need and work task process stage of each participant by asking about their chosen master thesis topic, their familiarity with the topic, and the intended use for found information. This was assumed to provide insight into the participant's context, which is emphasized as valuable by Spink et al. (1998, p. 617), and to mitigate potential label effects. The pre-search interview was designed as a classic reference interview, with a structure reflecting the five filters proposed by Taylor (1968).

One exception was filter 4: 'relationship of inquiry description to file organization' which was only involved during construction of query 2 in the pilot test, as participants were asked to search ERIC on their own. As will be shown in section 4.7, it was involved in constructing both queries in the final data collection procedure where both searches were mediated. Therefore, the added intention of the pre-search interview in the final data collection procedure was to arrive at a version of a Q4 need, which the participant found acceptable.

An important part of the participant-intermediary interaction took place between the two think-aloud interviews, both in the pilot and the final procedure. This interaction could be called a post-search interview as such, but it differs from the final post-search interview in that the focus was on constructing an expanded version of the first query. The aim was to use feedback from both the participant and the system in creating an expanded second query with synonyms and related terms, which utilized Boolean 'AND' and 'OR' where appropriate. The resulting expanded query retrieved the results assessed in the second think-aloud interview.

The post-search interview focused on participants' immediate reflections of the perceived usefulness, quality and complexity of the queries used. In the revised data collection procedure, questions regarding previous experiences with the use of Boolean operators were moved to the post-search interview to avoid bias in the participants evaluation of queries and results. For example, drawing on a previous negative or confusing experience with using Boolean operators in the pre-search interview prior to the think-aloud sessions, might had influenced how the queries and results were judged. See appendix C for the interview guide used in the pilot tests, and appendix E for final interview guide.

4.6 Pilot testing

Pilot testing consisted of complete search sessions with three different participants. The purpose of this extensive testing was, as mentioned, to detect any issues that needed be corrected or decided on concerning the study design, the experimental setting, data collection and quality of data collected. See Appendix B for the study protocol used in the pilot study. Great care was taken to explain the procedure of data collection to the participants before any data was recorded. The ambition was to observe students' natural

search behavior, query construction and interaction as it occurs in a real-world setting, albeit via a digital meeting service.

Participants were asked to search the ERIC database for research and sources pertaining to their master thesis project, and to do so in their own manner. They were encouraged to make refinements to their searches until satisfied with the results or were of the impression that searching further wouldn't produce more relevant results. The intended outcome of these instructions was that system generated relevance feedback and subsequent iterations ideally would enable the participants to construct a query that was as closely connected to their present information need as possible, reflecting potential dynamic properties of the need. Two of the pilot test participants had used the ERIC database before.

Using this final search as a starting point, the second search session began with the author and participant working together on expanding the last query from the previous session, using Boolean operators AND and OR, as well as truncation and phrases. The author suggested synonyms and related terms to be included, based on the previously observed searches in the first session. The participant was asked for confirmation as to whether these were familiar and considered appropriate or not. This was done to ensure that the query still represented the current information need of the participant to the greatest extent possible. The settings in terms of search mode and expanders within the EBSCO*host* interface were the same for each participant and each session to ensure similar test conditions throughout the pilot tests. Participants could however change these settings for example by limiting results to only peer reviewed content or adjust the publication year range, as mentioned.

During these pilot sessions, it became quite clear that two of the three participants experienced varying degrees of reluctance towards constructing an initial query. Comments like "Well, I know that you know this system way better than me, so I'm not quite sure what words are the right ones here…" signaled that constructing a query while being observed was demanding and causing uncertainty. The author assisted by answering questions about the interface or settings but tried to avoid influencing the construction of the initial queries. This proved quite challenging if not impossible, as participants asked for confirmation on the 'correctness' of the queries they constructed in terms of which operators to use, and how they should be combined. The test participants constructed 3, 2 and 6 queries respectively, in the first part of their session. It was the author's distinct impression that all three

participants experienced relief when the second, expanded query was constructed in collaboration.

Participants were encouraged to read abstracts or full texts if they felt it would aid in their assessment process. When it came to relevance assessments of the retrieved results, the impression was that this was done in a more hurried manner than perhaps was usual for all three participants for both the first and second search session. When asked about how he or she assessed a certain reference in the result list, one participant stated that "Normally, I would have read the abstract for this paper" but chose to scroll on to the next result after looking only at the title. This made it clear that the participant did not deem the situation 'normal' in the sense that time could be spent reading abstracts and larger excerpts of text. With regards to the interview approach however, the general impression was that few reminders or interventions were needed to elicit relevance statements and other considerations as such, from all three pilot test participants. Below, Table 2 gives an overview of the scope of the pilot test.

Participant	**Pilot participant A**		**Pilot participant B**		**Pilot participant C**	
Search session	S1	S2	S1	S2	S1	S2
Total no. of run queries	3	1	2	1	6	1
Total no. of results (Q = Query)	Q1: 5 Q2: 334 Q3: 5	Q1: 230	Q1: 220 Q2: 125	Q1: 146	Q1: 258 Q2: 25681 Q3: 4494 Q4: 45 Q5: 0 Q6: 7	Q1: 114
Time spent assessing results	12:38	9:36	5:51	5:32	44:52	22:34

Table 2: Overview of data collected in the pilot tests

The main learning point from the pilot tests was that asking participants to construct a query all on their own for the first part of the search session proved to be quite demanding for them. As the individual search experience varied quite significantly between the three participants, this also resulted in one participant constructing elaborate and advanced queries with confidence, while others were unsure of how they should proceed. Although they all managed to construct queries that yielded results they assessed as relevant, the construction process itself was interpreted as rather demanding for them, by the author. This was first and foremost considered to put an unnecessarily large and negative burden on participants. Second, it reduced the chances of setting the stage for an open and ideal think aloud interview interaction during relevance assessments.

Another important point was that conflicting roles on the author's part made it quite challenging to keep in line with the initial research design, which aimed at observing natural query formulation processes among the participants. This was especially the case when participants asked for help in constructing initial queries, or experienced difficulties interpreting what had gone wrong if a query retrieved very few or no results. This resulted in an internal conflict regarding the role of being a university librarian and a master's student with a personal research agenda at the same time. In other words, the goal of observing un-mediated information behavior conflicted with the mediator role of helping the participants in using and making sense of the IR system they were interacting with.

Designing an experimental procedure that allowed for user-constructed queries proved difficult. Judging by experiences in the pilot test, it would in any case not accommodate the capture of natural search behavior amongst the participants. Excluding a 'natural' query construction process for the initial searches eliminated the possibility of observing what tactics the participants might have used when facing difficulties in query construction and reformulation, which would have been an interesting aspect to examine.

This exclusion did however suggest a data collection procedure involving the mediated construction of versions of Q4 needs in an interactive search setting, which was also considered very interesting. At the same time, it made possible a user-centered study of Boolean queries including and excluding certain operators. This change in the research design still aimed at involving the tripartite user-intermediary-system interaction described in Ingwersen (1996, p. 17) to a large degree, through which the user's information need is

considered as reformulated due to interaction with the retrieved information objects in the system, and involving inquiries from the intermediary as to the relevance of retrieved objects.

Reflecting on the initial research questions for this thesis in light of the pilot tests, it became clear that the perceived usefulness of Boolean queries could rather be inferred by narrowing the scope of data collection to focus mainly on information needs and the mediated construction of queries, and relevance assessments of results. In addition, a choice was made to investigate the perceived usefulness of queries using only 'AND', and queries using 'AND and 'OR'. This choice was inspired by the research questions investigated by Lowe et al. (2020). The first and second pilot test interview data were transcribed, but subsequent analysis was not prioritized as time had to be spent on developing revised data collection procedures.

4.7 Final research design

As a result of lessons learned in the pilot tests, the research questions were reformulated into those presented in the introduction of the thesis. Mason's (2018, p. 26) table is repeated here in the context of the final research questions:

Research questions	**Data sources and methods**	**What could they yield?**
1. How do graduate students within the educational sciences perceive the usefulness of simple and advanced Boolean queries, when searching a subject database?	Graduate students: Concurrent think-aloud, and pre- and post-search interviews Graduate students: Concurrent think-aloud interviews	Statements from all three interview types can shed light on the perceived usefulness of simple and one advanced Boolean query during an IR situation. Think-aloud statements narrow the scope to information need dynamics during searching. Relevance assessments of results from one search session using a simple Boolean query ('AND') and one session using an advanced

		Boolean query ('AND', 'OR', truncation, wildcard) show what criteria the students rely on, adding depth to the understanding of these processes.
	Graduate students: Screen recording of two search sessions	Number and rank of viewed results for each search. Number and rank of assessed results for each search.
2. How is the perceived usefulness of simple and advanced Boolean queries among students related to the students' information needs and relevance criteria?	Pre-search interview transcriptions: Inductive analysis	Descriptions of students' information needs, and their context give insight into information need type, making it possible to examine the relations between perceived usefulness of queries, and information needs.
	Think-aloud interview transcriptions: Deductive analysis	Frequency of assessments pertaining to relevance criteria
	Post-search interview transcriptions: Inductive analysis	Descriptions of students' experiences with using Boolean operators in searches

Table 3: Final research questions and proposed data sources and methods

To sum up, the following changes were made to the research design: To accommodate for the expected reluctance in constructing an initial query amongst participants, data gathered in the online recruitment form was used as a basis for constructing draft versions of one simple and one advanced Boolean query for each participant, before each search session. A

post-search interview was included to collect statements about the participants' perceptions of the two queries, and previous experiences with using Boolean operators.

4.7.1 Construction of the simple and advanced queries

The decision was made to construct draft versions of a simple Boolean query using only the AND operator, and then expand this simple query into a more advanced version, using the OR operator and truncation. The purpose of this expansion was to include synonyms and related terms to those used in the first query, as well as single and plural form of nouns. An increased number of potentially relevant results was the expected outcome of the advanced query. The ERIC thesaurus and retrieved results of test queries were studied to identify terms that could be used in the queries. The ERIC collection was searched using different variants of queries until results lists seemed to be returning results that could be deemed relevant, with regards to the information submitted in the online form by each participant.

As can be seen later in section 5.2, both the simple and advanced version of queries for both participants utilize exact phrases to an extensive degree. The reason is that quite a few of the concepts that were stated as of interest by the participants in the online recruitment form, can be considered multiple-word concepts that for example specify certain *kinds* of learning, i.e., "play based learning", and "language learning". The use of phrases in quotation marks in a query in the EBSCO*host* interface enables these concepts to be searched as exact phrases, rather than the default near 5 (N5) proximity operator returning any results where the words in the query are 5 or less words apart in any direction (EBSCOhost, n.d.-b). All multiple-word concepts were searched as exact phrases in both the simple and advanced queries, to provide the same level of specification of such terms across queries and participants. The aim was not to explore the perceived usefulness of the use and non-use of exact phrases in queries as such, but rather the use and non-use of truncation and Boolean OR. Wildcard was used in one of the advanced queries.

A parallel to this process of interpreting the content in the online forms, is seen in the way intermediary query formulation is put into the context of Ingwersen's (1996) cognitive IR model, by Sormunen (2000):

> Query formulation is a process where an intermediary perceives and interprets a user request and translates it into a query. Two basic transformations take place: from the

> linguistic level (the user request in natural language) to the cognitive level (intermediary's knowledge structure) and back to the linguistic level (Boolean query). The way that an intermediary interprets a user's request depends on his/her current cognitive structures. (Sormunen, 2000, p. 32)

The user request in the online form was however not stated in natural language as such, but composed of the responses given to these two points:

- A brief written description of what kind of research they are searching for at the present time, in relation to their master thesis
- A query using terms of their own choice, as they would have planned to use, when searching the subject database ERIC

Before going further, an important point of reflexivity is due here. In constructing the draft versions of queries, a range of relevance assessments were necessarily carried out by the author. Relevance was assessed in relation to the information provided by the participants in the submitted online forms. This was done in an intellectual topicality sense, in that results were judged as to the presence of topics, their aboutness, to determine which terms were to be used in the queries. This highlights some of the very interesting challenges that arose in developing the revised research design: What elements should a user-centered study of Boolean queries contain, and how do these elements relate to a user- and system-oriented research perspective?

Realism is represented by the participants' own relevance judgements, and by using their individual information needs and work task situations as starting points. Experimental control is attempted in constructing queries with and without certain Boolean operators and truncations, exact phrases for all multiple-word concepts, as well as using the same IR system and settings for all participants. Two interesting challenges arose from these choices: If a participant's information need is to be considered as potentially dynamic within a single IR session, the simple and an advanced versions of queries should reflect this inherent dynamism. Second, the characteristics of the queries with regards to number of terms and operators, might be considered unrealistic compared with participants regular query construction. Each of these challenges will be subjects for discussion in section 6.

4.8 Final data collection procedure

Data collection sessions followed the procedure described in Appendix D. After the conclusion of each pre-search interview, the prepared draft version of the simple query was presented on the shared screen via Zoom and explained to the participant. The purpose of the presentation was to confirm the participant's present information need and make changes to the query if needed. During this presentation and explanation of the first query, extra focus was kept on the participant's response, and whether it resonated with the answers given in pre-search interview. It was considered important to emphasize to the participant that the query was constructed based on the author's assumed appropriateness of search terms. Statements on the author's part like "I carried out some searches to see what terms might be used to describe these concepts, but I am still a bit unsure", were intended to invite the participant to comment, ask about and perhaps also correct terms according to their experiences with the topic, and information need, thereby involving Taylor's (1968) fourth filter**.**

A draft version of an advanced query had also been prepared before the data collection session and was presented and explained after the assessment of the results of the simple query was complete. This is where perhaps the most challenging aspect of the final research design emerged; While the work task of the participant still could be considered stable, the aim was to reflect potential changes in information needs that might had occurred as a result of relevance assessments and interaction in the first search, in construction of the advanced version of the query. The chosen strategy was to again ask the participant to comment on the appropriateness and discuss any terms or concepts that the participant had encountered and commented upon during relevance assessment. The modifications of advanced queries are described in detail in section 5.1.

The sequence of the search tasks was not rotated across participants, although such rotations might be expected for similar IIR experiments to counter for learning and order effects (Kelly, 2009, p. 50). This is because the advanced Boolean query was intended to function as an expansion of the first, simple Boolean version. In other words, the sequence of a simple query followed by an advanced elaborated version was intended to reflect an assumed strategy in an IR-situation: The initial query is expanded to increase the number of potentially relevant results, using the first query as a basis for this expansion. A reverse

order of search tasks would therefore not make sense. One effect that ideally would be controlled is that of fatigue, and the possibility that participants were closer to being tired of the experiment during the second, advanced search.

To provide clues to answering RQ1, it was important to get participants to reflect on the perceived usefulness of the queries in question, and for RQ2, to see these answers in the context of relevance judgements made in accordance with the reported information need. In addition, the data collection aimed at gaining insight into the participants' general experiences with using Boolean operators. While the participants' responses to the queries and results yielded a rich set of qualitative data, they still represent users' experiences in a single point in time. Questions in the post-search interview about previous experiences and incidents supplement this to a certain extent.

4.9 Resulting data and preparation for analysis

The resulting data from each search session consisted of:

1. sound recordings of pre- and post-search interviews, and think aloud interviews for both searches, stored for playback on the Nettskjema server

To prepare the interview data for analysis, all sound recordings were transcribed in Word in a table format including time codes for verbalizations. This allowed for export to NVivo, where the transcripts were synchronized for playback with the screen recordings. All transcripts were controlled against the sound recordings for accuracy, as advised by Braun and Clarke (2006, p. 88).

2. Screen recordings of the two searches

Screen recordings were saved in Screencast-O-Matic, then exported as .AVI-files to ensure compatibility with processing in NVivo and synchronized playback with interview transcripts

3. Printed search history showing query and number of results for both searches

Search histories were saved as pdf-files, providing descriptive data for each search

4. Export of a ranked list of references for each search, reflecting the order presented in the result list

A complete result lists for each query was exported using the EBSCO Export Manager. This RIS-file was then saved in a separate EndNote library and exported to Excel, which allowed for an efficient way of visualizing duplicate results across the two queries, and their respective rank in each list viewed side-by-side. These ordered result lists were also used to get an overview over which results from the two queries were verbally assessed.

From here on, the elements of the data collection are referred to in the following manner, to aid with readability: The search session in which the results of the simple query were being assessed is sometimes abbreviated as SQS (Simple Query Session), the advanced as AQS (Advanced Query Session). Participants are abbreviated as P1 up through P7.

4.10 Analysis of interview data

Both a thematic analysis and content analysis approach to the interview data were considered fruitful. Content analysis of qualitative data is similar to thematic analysis, but differs in that it often lends itself more easily to quantification by focusing on details and providing for example for the frequency of occurring terms (Braun & Clarke, 2006, p. 98). Hsieh and Shannon (2005, p. 1278) define qualitative content analysis as "[...] a research method for the subjective interpretation of the content of text data through the systematic classification process of coding and identifying themes or patterns". How the codes and themes are generated will differ according to the chosen approach, with 'directed content analysis' using established theory or existing research findings as the basis for categorization (Hsieh & Shannon, 2005, pp. 1281-1283). This is in contrast to a more inductive 'conventional content analysis' where codes and themes emerge through repeated reading and interpretation of the textual data (Hsieh & Shannon, 2005, p. 1279).

For RQ1[6], the goal was to learn more about how the participants perceived the usefulness of the two queries used in the search sessions. Part of RQ2[7] was aimed at gaining insight into the participants' information needs. An inductive conventional approach generating descriptions and explanations in context was chosen for these two areas of interest, by

[6] RQ1: How do graduate students within the educational sciences perceive the usefulness of simple and advanced Boolean search strategies, when searching a subject database?

[7] RQ2: How is the perceived usefulness of simple and advanced Boolean search strategies among students related to the students' information needs[...]?

allowing for a wide variety of themes. The pre- and post-search interviews were coded in NVivo using this inductive approach.

A more deductive, directed approach was deemed appropriate in terms of providing data to answer RQ2[8], which aimed at describing how information needs, relevance assessments, and the perceived usefulness of the queries might be related. The think aloud-interview transcripts were coded in NVivo, using Saracevic's (2017, pp. 57-58) groups of clues and criteria based on object characteristics (content, object, validity) and human characteristics (use or situational match, cognitive match, affective match, belief match). This was done to elicit and describe quantitatively what observed relevance criteria the participants relied on during assessments. How each group is defined and understood in the coding process is explained in section 4.10.3. It is important to note that this directed approach still allowed for new codes and themes to emerge from the data. Rather than limiting the analytical scope, it was intended to provide an established research-based platform on which to build the analysis. In sum, a combination of both a quantitative and a qualitative approach to content analysis was chosen, as suggested by Zhang and Wildemuth (2017, p. 319).

4.10.1 Background for the deductive analytical approach

The choice of Saracevic's human and object groups of clues and criteria (2017, pp. 57-58) represents an expected 'range of variation' as defined by Kelly (2009, p. 39), in that the generalized groups are considered to encompass the range of criteria that can be anticipated in the data. The degree of exhaustiveness and exclusiveness of these categories was however an open question at the start of the coding process, which is described in section 4.10.2.

The deductive analytical approach aims at examining the data in this study with the relevance dimension of *criteria* theoretical context, expressed through the Saracevic's groups. Frequency of stated relevance criteria are considered as a ratio level of measurement as defined by Kelly (2009, p. 43), as it expresses the number of incidents of what is measured. The frequency is the result of interpretations on the authors part, made through the coding process described in the subsections below. Ingwersen (1996, p. 17)

[8] RQ2: How is the perceived usefulness of simple and advanced Boolean search strategies among students related to the students' information needs and relevance assessments?

reviewed earlier research investigating what relevance criteria users rely on, stating that "cognitively speaking, we may assume that the variety in relevance assessment categories empirically extracted from users [...] corresponds to the variety of the formation states or foci of the information need evolution". This assumed correspondence is taken as a point of departure for the analysis and discussion of the empirical data in terms of stage in task and characteristics of information needs and relevance assessments. It is considered to provide an interesting context for understanding the participants' perceived usefulness of Boolean queries.

The interdependence and dynamic interplay between 'object' and 'human' groups of criteria suggested by Saracevic (2017, p. 58) offers an understanding of relevance assessments as seen from the user's perspective. Frequencies of occurrence and co-occurrence of criteria within the two groups are chosen as measures because they are anticipated to give insight into these processes.

4.10.2 Development of the deductive coding procedure

A preliminary, directed coding was performed on the think-aloud transcript from the simple query session of the first participant (P1). It became clear that some of the relevance assessments made could be considered latent in nature, in the way that they were not explicitly verbalized by the participant. A typical example of such assessments would be the participant stating, "Number eleven seems interesting", while viewing the result list, which included title, author, publication year and subject terms in the presentation of the result in question. The interview situation did not always allow for a natural way of following up with questions regarding what exact criteria the participant relied on. Without any more information about this assessment, it could not be determined what object-related criteria were used by the participant. During data collection, efforts were made on the authors part to follow up as many of these latent statements during interviews, as possible. In the cases where the participant read parts of titles, subject terms or abstracts out loud, or the criteria for assessment was unambiguous, verbalizations were coded to the respective category.

Hence, the final deductive coding procedure followed this understanding of verbalized criteria, coding only those criteria that were explicitly stated and those that where conclusive in terms of content read out loud, resulting in transcribed content. A second directed coding was then performed with this procedure, on the think-aloud transcripts of

both search sessions for P1. This led to an established procedure in NVivo which was then used for the remaining think-aloud transcripts. Along the way, the coding process revealed several new criteria within the established hierarchy of categories and groups, that were added to the final coding scheme as shown in section 4.10.3.

Time did not allow a complete second coding from scratch for all think-aloud transcripts, as would be ideal for establishing intra-coder reliability and test the coding structure for robustness. However, all highlighted (coded) content in NVivo was reviewed 10 days after the first complete round of coding to get a second look at the decisions made during the initial procedure. This led to the recoding of about 50 verbalizations across all transcripts, and the un-coding from relevance criteria codes of 25 verbalizations. The large amount of uncoded content was mainly due to a manual mistake when a large section of one of the transcripts (19 consecutive verbalizations) was coded to a single category during the first round.

4.10.3 Groups of relevance criteria

The aforementioned generalized groups of criteria identified by Saracevic (2017, pp. 57-58) provided the initial structure and hierarchy in the deductive coding. Although the scope of each group is described to a certain extent by Saracevic through the listing of examples, the coding process was anticipated to result in an adapted and developed set. In other words, the structure provided a conceptual starting point as to the grouping of criteria, but not necessarily individual criteria themselves.

An important point to stress here, is that Saracevic's phrasing of the 'Human characteristics' groups as different kinds of matches (situational match, cognitive match, affective match, belief match) is not equated as an assessment of an object as *relevant*. It is rather interpreted as a match in the sense that criteria pertaining to that group are activated, and that they may well be so when a result is assessed as *not relevant*.

Table 4 below shows how the initial descriptions were understood and operationalized in the coding procedure:

Object characteristics	
Groups	**Criteria** (description)
Content	**Topic** (The topic or subject of the document) **Quality** (The perceived quality of the document. In the coding procedure, 'Quality' as a criterion is operationalized as belonging to the 'Validity' group of criteria, and covered by the criteria in that group) **Depth** was combined with **Scope:** (The extent of treatment or depth/focus in a study. Can be in terms of age of population in study, a span of years for a set of examined documents in a study or similar) **Currency** (The publication date or year of a document) **Treatment** (They way in which the topic or subject of a document is being discussed, perspective on topic as expressed in document) **Clarity** (not included as a content criterion, but understood as strongly linked to the cognitive match criterion 'understanding')
Object	**Type, Organization, Representation, Format, Availability, Accessibility, Costs** (These were not included as single criteria, but instead operationalized through the identified and added object criteria described in the next table)
Validity	**Accuracy of information provided, Authority, Trustworthiness of sources, Verifiability, Reliability** (These were not included as single criteria, but instead operationalized through the identified and added validity criteria described in the next table)
Human characteristics	
Groups	**Criteria** (description)
Use or situational match	**Appropriateness to situation, or tasks** (An assessment of the object as to the appropriateness in terms of the participant's work task) **Usability** (Not included as a criterion, but instead operationalized through the other criteria in this group) **Urgency** (An assessment of the object as urgent in terms of usefulness for the participant) **Value in use** (An assessment of the object in terms of its value in the context of the participant's work task)
Cognitive match	**Understanding** (Assessment based on understanding) **Novelty** (An information object is assessed as new or original) **Mental effort** (Not included as a criterion)

	Link to previous knowledge (An information object is assessed based on participant's existing knowledge and is explicitly linked to this knowledge)
Affective match:	**Emotional responses to information** (Not included as single criteria, but instead operationalized through the ones listed by Saracevic and added criteria described in the next table) **Fun** (Fun as an emotional response to the assessment of an information object) **Frustration** (Not included as a criterion) **Uncertainty** (Uncertainty as an emotional response to the assessment of an information object)
Belief match:	**Personal credence given to information** (Operationalized as it is described) **Confidence** (Not included as a criterion)

Table 4: Initial groups of criteria and descriptions

The developed set was in part constructed through the coding process itself, but was also aided by studying the criteria identified by Barry (1994), and the overview in Taylor (2013, pp. 543-544), especially in situations of doubt as to the correct categorization of certain verbalizations. Table 5 shows initial groups from Saracevic (2017, pp. 57-58), what new criteria were added and how these were defined for the purposes of coding.

Human characteristics	
Initial groups	**Added criteria** (description)
Content: "topic, quality, depth, scope, currency, treatment, clarity"	**Geographic location of study** (The geographic location of the study reported in the object being assessed) **Certain terms (search via ctrl+f) occurring in full text** (Whether or not a certain term is present in the full text document when searched by the participant using ctrl+f) **Certain terms (search via ctrl+r) occurring in result list** (Whether or not a certain term is present in the result list when searched by the participant using ctrl+f.) **Methodology** (Type of study or choice of methodology) **Conclusions** (The conclusion section of a full text document) **Sub criteria of Currency**: **Currency not important** (The publication date or year stated as not important by a participant) **Findings** (The findings or results section of a full text document) **Sources referenced** (The sources referenced in the document)

Object: "characteristics of information objects, e.g., type, organization, representation, format, availability, accessibility, costs"	**Abstract** (The abstract section of a document or bibliographic reference) Subcategory of Abstract: **Actively not reading abstract** (The abstract stated as not important/of little utility by a participant) **Full text** (The full text version of a document) **Full text language** (The language in a full text version of a document) **Language of bibliographic reference** (The language of bibliographic reference) **Rank in result list** (The rank of a certain result within a result list) **Subject terms** (The subject terms used to describe a result) **Title (**The title of a document) Subcategory of Title: **Length of title** (The number of words in a title) Subcategory of Title: **Recognizes title from simple query** (The participant recognizes the title of a result from the previous, simple query)
Validity: "accuracy of information provided, authority, trustworthiness of sources, verifiability, reliability"	**Author** (The author or authors of a document) **Journal** (The journal of a document) **Peer-Review** (Whether or not the document is peer reviewed) **Publisher** (The publisher of a document or journal)
Human characteristics	
Initial groups	**Added criteria** (description)
Use or situational match: "appropriateness to situation, or tasks, usability, urgency; value in use"	Criteria listed to this group in Saracevic (2017, pp. 57-58) and described in the previous table were exhaustive for coded statements
Cognitive match: "understanding, novelty, mental effort. Link to previous knowledge"	Criteria listed to this group in Saracevic (2017, pp. 57-58) and described in the previous table were exhaustive for coded statements
Affective match: "emotional responses to information, fun, frustration, uncertainty"	**Curiosity** (Participant stating or signifying that he/she became curious as a result of assessment) **Enthusiasm** (Enthusiastic response during assessment) **Empathy** (Participant expressing emphatic concern)
Belief match: "personal credence given to information, confidence"	Criteria listed to this group in Saracevic (2017, pp. 57-58) and described in the previous table were exhaustive for coded statements

Table 5: Criteria groups from Saracevic (2017, pp. 57-58), and added criteria

During the review of the initial coding, statements that had remained in the groups 'Use or situational match', 'Cognitive match', 'Affective match' and 'Belief match' after the first round were all coded to single criteria within the group to add precision in the analysis. One new code ('Empathy' as a subcategory to 'Affective match') was added during the review. None of the initial codes were removed. During later data analysis, the criteria 'Title recognized from simple query' was added as a subcode to 'Title' to be able to detect the number of results in this category.

Description of 'anchor examples' of statements coded to groups are recommended by Assarroudi et al. (2018) are provided below in Table 6. Due to space considerations, examples are provided here for verbalizations that were coded to multiple categories, and for one that generated doubt as to categorization. A full version of the codebook used in NVivo with descriptions of each code can be found in Appendix G.

Anchor example (participant, think-aloud session)	Criteria	Rationale and comment
Right, semi structured interviews, that's an approach I'm also considering using... (P4, SQS)	**Object characteristics:** (Content) Methodology (Object) Abstract **Human characteristics:** (Use or situational match) Value in use	Participant inferred the methodological approach in a document by reading abstract, and stating that this related to the task at hand (planning the data collection procedure in the thesis)
But that one, «towards holistic supporting of play based learning", I'm thinking that is relevant both ways, cause that "holistic" way of thinking, that corresponds directly to the way I'm thinking (P1, SQS)	**Object characteristics:** (Object) Title (Content) Topic (Content) Treatment **Human characteristics:** (Cognitive match) Understanding	Participant read excerpt from title out loud, stating that topic (play-based learning) was relevant, and that the term 'holistic' was recognized as a familiar way of treating the topic Comment: Generated doubt as to categorization into topic or treatment. Conclusion was that both were considered appropriate for this verbalization.

Student perspective on low English... proficiency in Oman, not relevant either, both not what I am looking for, and not an area that is comparable" (P2, AQS)	**Object characteristics:** (Object) Title (Content) Topic (Content) Geographic location of study **Human characteristics:** Appropriateness to situation, or task	Participant read title, stating that the result was not on topic, and that the location of the study rendered it inappropriate to the task.

Table 6: Anchor examples of coded statements

The 'Matrix Coding Query' function in NVivo was used to export the number of verbalizations coded to each criterion within each think-aloud transcript. This query identifies references (coded content) assigned to designated codes and produces a tabular overview with codes and number of coded instances for each file (transcript). These numbers where then exported to Excel for combination across participants and calculations of totals and percentages. As can be seen in the examples shown above, each verbalization was potentially coded to several categories, depending on the interpretation of the content, and chosen operationalization of criteria.

A goal was to also code each assessed result either as relevant, partially relevant, or not relevant based on the participants' verbalizations. However, the final data collection procedure failed to accommodate for such grading in a concise manner, as the degree of relevance of each result was not explicitly asked for by the author. A three-level grading of non-relevant/partially relevant/relevant might have been possible to ascertain from transcripts, but after some consideration and given the limits of the timeframe of this project, the decision was made to focus on the analysis of criteria.

5 Results

Participants were graduate educational science students in the planning stage of their master thesis, who all volunteered to contribute to the project. Data collection that generated the results reported here, took place between March and September 2021. Data and queries from the SQS think-aloud interviews are shown in columns with blue headings, and from AQS think-aloud interviews in yellow.

5.1 Scope and amount of data

A selection of descriptive statistics is provided here to give a coherent sense of scope and amount of data gathered, and to aid the presentation of additional results in section 5 and subsequent discussion in section 6. Table 7 shows an overview of participants and duration in minutes of the segments in each of the data collection sessions:

Participant	Pre-search interview	SQS Think-aloud	AQS Think-aloud	Post-search interview
P1	7	25	19	3
P2	7	6	28	12
P3	6	12	8	7
P4	12	37	9	6
P5	5	8	9	4
P6	13	14	7	5
P7	3	5	5	4

Table 7: Time spent during data collection sessions

Each entire data collection session (pre-search interview, explanation of SQS query, SQS think-aloud, explanation of AQS query, AQS think-aloud, post-search interview) lasted about one hour (57 minutes) on average, per participant. The sessions generated a total of about 6,5 hours of sound recordings that were transcribed in preparation for analysis. The think-aloud interviews represented about 3 hours of the material. To further illuminate the volume of data processed and collected, a total of 149 search results were relevance

assessed by the participants, which lead to relevance verbalizations that were coded to a total of 651 occurring criteria.

The participants spent an average of 15 minutes assessing the results in the SQS, and 12 minutes on the results in the AQS. There were however substantial differences in time spent among participants, varying between 5 and 37 minutes (median= 12 minutes) for the SQS, and 5 and 28 minutes (median= 9 minutes) for the AQS. When looking at individual search sessions, participants either spent slightly equal or less time assessing the results in the AQS as they did when assessing the results in the SQS, with one exception for participant 2 (P2). This exception probably occurred because the simple query prepared for P2 only returned three results. Large variation across the relatively small number of participants makes it difficult to detect any obvious tendency in terms of time spent.

Some of the other SQS queries also returned relatively few results, and the majority returned less than 31 results. In these five, participants scrolled through the entire result list. For the two remaining larger sets in the SQS, P3 scrolled until result 12 of 441, and P5 scrolled until result 20 of 576. Table 6 shows the number of retrieved results for each of the two queries for each participant, and at which point in the result lists each participant stopped looking further down in the list:

	Simple Query		**Advanced Query**	
Participant	Retrieved results	Viewed until	Retrieved results	Viewed until
P1	31	31	52	52
P2	3	3	29	29
P3	441	12	561	29
P4	14	14	72	50
P5	576	20	5711	29
P6	18	18	38	28
P7	7	7	230	26

Table 8: Overview of retrieved results in ERIC, and no. of viewed results

The queries used in the AQS retrieved more results than the simple ones, which was expected as they were intended to be expansions of the initial queries. As can be seen, the tendency was for participants to stop scrolling in the list at around result no. 29 in the

lists of the advanced queries that had remaining results. The exceptions were P1 and P4, who viewed the entire list, and the entire first page of results, respectively.

A central source of data are the verbal assessments in the think-aloud interviews, which were coded to relevance criteria groups in NVivo. Table 5 shows the number of results for each of the two queries for each participant, and how many of the results that were verbally assessed.

	Simple Query		**Advanced Query**	
Participant	Retrieved results	How many of results assessed verbally	Retrieved results	How many of results assessed verbally
P1	31	19	52	36
P2	3	3	29	25
P3	441	4	561	8
P4	14	5	72	8
P5	576	2	5711	4
P6	18	10	38	9
P7	7	5	230	11
Total		**48**		**101**

Table 9: Overview of retrieved results in ERIC, and no. of verbally assessed results

As shown, the number of verbally assessed results is more than double in the AQS than in the SQS, even though participants on average spent less time assessing these results. Part of the explanation for this is that 21 of the verbally assessed results in the AQS had already been seen in the SQS. This was a fact which the participants in question would quickly acknowledge verbally before moving on in the list. These verbalizations were counted and coded as titles that were recognized from the SQS by the participant.

There was however a total of 17 results that occurred in the SQS result lists, which were only assessed during the AQS. Table 6 shows the total number of verbally assessed results in the AQS, and how many of these were (i) already assessed in the SQS, (ii) present in the SQS but only assessed during the AQS, and (iii) verbally assessed results unique to the AQS:

	Advanced Query			
Participant	No. of verbally assessed results	(i) Already assessed in the simple query session	(ii) Results appearing in the simple query but only assessed during the advanced	(iii) Verbally assessed results unique to the advanced query
P1	36	14	7	15
P2	25	2	0	23
P3	8	2	5	1
P4	8	2	1	5
P5	4	1	1	2
P6	9	0	3	6
P7	11	0	0	11
Total	**101**	21	17	63

Table 10: Distribution of assessed results in the AQS.

The majority (63) of the verbally assessed results in the AQS concern those that were unique to that session. As previously stated, a sum of 149 results were assessed verbally by the participants during the SQS (48) and AQS (101), of which eligible statements provided the data from which frequency of relevance criteria are obtained. The resulting frequency of relevance criteria is presented as combined for all participants in section 5.3. In the following, each participant is presented along with the interpreted information need and stated intended use for found information, together with the queries used in each session.

5.2 The participants

5.2.1 Participant 1 (P1)

P1's master thesis topic was the role of play in the early school years and how it can help with the transition from kindergarten to primary school. P1 felt strongly that this is an important concept for teachers to recognize, and that published research findings on the topic might not be followed up in practice. It had been a personal topic of interest for P1 for several years. P1 had looked at the ERIC database during library instruction but had not used it for searching, preferring instead to use the Oria discovery tool. During the last semester,

several fellow students had provided P1 with useful sources on the topic during workshops and in other similar contexts. The last time the topic had been searched by P1 was about a month prior to data collection. P1 was still looking for definitions on the concept of 'play', and examples in peer reviewed research of how play can be implemented in a school setting. As a result of the pre-search interview, the current information need and intended use of found information for P1 were interpreted in this way by the author:

Information need	**Intended use**
Definitions of the concept of «play», and peer reviewed research containing examples of play-based learning in primary education (grades 1 and 2).	To gain overview and knowledge on chosen research theme before defining a research question for the master thesis.

P1's information need was considered as a variable conscious topical (CIN) type need, in that the participant was looking for further information in a subject matter that was stated as well known. Using information provided by P1 in the previously submitted online form, the simple query presented in the table below was constructed with two major facets representing the concepts considered involved. It been prepared in a Word document before the data collection session and was presented and explained to P1 via shared screen directly after the pre-search interview. P1 confirmed that the terms used were familiar, and that their combination seemed appropriate given the present information need. Therefore, no changes were made to the proposed query.

P1 Simple Boolean query	
"play based learning" AND "primary education"	31 results (peer review filter activated by participant) 19 results assessed verbally 7 abstracts viewed 0 full texts viewed

The author pasted the simple query into the advanced search screen of the EBSCO*host* ERIC interface. Control over the shared screen was then given to the participant using the 'remote control' function in Zoom, upon which P1 could interact with the database by clicking the

search button and viewing the retrieved results list. This procedure of inserting each query and then giving control via Zoom before each query was run was the same in all data collection sessions.

The peer review filter in the EBSCO*host* interface was activated after a few minutes by P1, who stated that faculty had emphasized the importance of peer reviewed research as source material. 19 results were assessed verbally during the SQS, of which 7 abstracts and 0 full texts were viewed.

A draft version of the advanced query was constructed before the data collection session, with the aim of including alternate spelling variations of 'play-based learning', and synonyms and related terms to 'primary education'. P1 was presented with this query, and confirmed that these additional terms were familiar, and still in line with the information need. As a result, no changes were made to the proposed AQS query either.

P1 Advanced Boolean query	
("play based learning" or "play-based learning" OR "learn* through play*) AND ("primary education" OR "elementary education" OR "primary school*" OR "elementary school*")	52 results (peer review filter activated by participant) 36 results assessed verbally 6 abstracts viewed 0 full texts viewed

36 results were assessed verbally during the AQS, of which 6 abstracts and 0 full texts were viewed. Three of these abstracts belonged to results that were assessed previously in the SQS, of which two abstracts had also been viewed during the first session (only the title belonging to the third abstract had been assessed in the SQS).

5.2.2 Participant 2 (P2)

P2 wanted to investigate how informal learning of the English language happens outside the classroom and how teachers can accommodate informal learning and utilize it as a resource during classes. P2 described having a general level of knowledge on the topic from own experience, through research literature, and through discussing it with adolescents who had explained how their informal language learning was taking place. P2 was looking for

published research or literature describing the actual phenomenon thereby providing a foundation, which could then be used to position the thesis within the research field. P2 was at an early point in the master thesis project and stated that only few results (2-3) were needed to fulfill the current information need, as P2 would usually print and read a few selected results, as a habit. P2 had had a quick look at the ERIC database (original version) in preparation for the data collection and stated that the previous search on the topic was conducted quite some time ago, in the preceding semester. The current information need and intended use of found information for P2 were interpreted in this way:

Information need	**Intended use**
Research that shows how adolescents learn English outside the classroom setting in their spare time, and examples of how this is integrated by teachers in formal educational settings.	To position participant's master thesis in context of earlier research

The information need of P2 was also understood as a variable CIN type need of by the author, as the participant stated that the topic was considered as being relatively familiar. The following proposed query was constructed with three major facets. It was presented and explained to P2, who stated that no changes seemed necessary:

P2 Simple Boolean query	
"language learning" AND "spare time" AND youth	3 results 3 results assessed verbally 1 abstract viewed 1 full text viewed

As previously mentioned, the simple query returned quite few results. All three were assessed, one of which abstract and full text were viewed. Accordingly, this session was quite short in duration. A draft version of the advanced query was constructed with the following aims: To include alternate spelling variations, synonyms, and related terms of 'language learning'. Second, to include terms that incorporated the autonomous learning aspect, and third, to include synonyms and related terms to 'youth' in single and plural form.

When presented with the query shown below, P2 stated that the additional terms were familiar, and that the query was still in line with the information need. No changes were made to the proposed AQS query.

P2 Advanced Boolean query	
("Language learning" OR "Second language learning" OR "Foreign language learning" OR EFL OR «english as a foreign language») AND ("Outside the classroom" OR "Self-regulated learning" OR "Autonomous learning" OR "Leisure time" OR "Spare time") AND (adolescen* or youth or teenage*)	29 results 25 results assessed verbally 7 abstracts viewed 1 full text viewed

25 results were assessed verbally during the AQS, of which 7 abstracts and 1 full text document were viewed. All abstracts and the full text viewed belonged to results that were unique to the AQS query.

5.2.3 Participant 3 (P3)

The topic for P3's thesis was the concept of life skills and inter-personal relations in a school setting, and how these issues affect the learning situation for students. P3 had knowledge on the subject from a previous university course and having read the course literature. Additional sources such as peer-reviewed research, books and published bachelor theses, master theses, and PhD theses on the topic had been consulted. P3 was looking for additional research and literature to gain more knowledge on the topic from different perspectives and stated that such literature would be relevant both for the master thesis, and an ongoing university course. P3 was not familiar with the ERIC database but reported having used Oria extensively. As a result of the pre-search interview, the current information need and intended use of found information for P3 were interpreted in this way:

Information need	**Intended use**
Research or literature on the topic of students' life skills, and on students' perspectives on mental health in school, limited to the senior	Gain knowledge on the topic, and utilize in master thesis and a course running parallel to the master thesis course in the fall semester

years of primary education (7th grade in the Norwegian school system)	

This information need represented another variable CIN in the author's opinion, as P3 reported having studied the topic. The query below was presented and explained to P3 who stated that the terms used were familiar and appropriate given the present information need.

P3 Simple Boolean query	
"life skills" AND "primary education"	441 results 4 results assessed verbally 2 abstracts viewed 1 full text viewed

4 of the retrieved results in the SQS were assessed, of which 2 abstracts and 1 full text document. The draft version of the advanced query included 'psychosocial skills' as a related term to 'life skills', and synonyms and related terms to 'primary education'. P3 added the combined term 'daily living skills' to the AQS query after noticing that it was used as a subject term in the first result in the SQS.

P3 Advanced Boolean query	
("life skills" OR "psychosocial skills" OR «daily living skills») AND ("primary education" OR "elementary education" OR "primary school*" OR "elementary school*")	561 results 8 results assessed verbally 1 abstract viewed 0 full text viewed

8 results were assessed verbally during the AQS, of which 1 abstract and 0 full text document were viewed. The abstract belonged to a result that was also present in the SQS, but only assessed in the AQS.

5.2.4 Participant 4 (P4)

P4 was interested in minority students' attitudes and perceptions towards their own native or first language, and how these might affect the students' learning situation when learning

English. For the planned master thesis project, P4 wanted to investigate students' experiences and how teachers handle the phenomenon in a classroom setting. The subject had also been proposed as a focus of investigation by faculty at the university. P4 had read literature related to the subject as part of completing previous courses and was currently looking for any research describing the use of first language in the classroom, as well as research examining minority students' attitudes towards their own native or first language. The primary purpose of locating this research was to identify potential research gaps. P4 had experience using the ERIC database (original version). Below is the information need and intended use of found information for P4:

Information need	**Intended use**
Research on the topic of minority students and their attitudes towards their own first language. How minority students' first language can be utilized by teachers in the English classroom.	To position participant's master thesis and add precision to research questions and direction, in context of earlier research

P4's need is considered as representing a variable CIN, owing to the participant's stated previous knowledge. The proposed simple query presented to P4 is shown below. No changes were made.

P4 Simple Boolean query	
"ethnic minority students" AND "first language"	14 results 5 results assessed verbally 3 abstracts viewed 2 full texts viewed

5 of the retrieved results in the SQS were assessed, of which 3 abstracts and 2 full text documents. After assessing the results from the SQS, P4 wanted to see what results a query using 'multilingualism' instead of 'first language' would retrieve. This query returned 0 results which triggered an interesting dialogue with regards to what related terms could be considered useful in subsequent searches. The draft version of the advanced query included synonyms and related terms to 'ethnic minority students' with truncation, and synonyms and related terms to 'first language'. No changes were made to the proposed AQS query.

P4 Advanced Boolean query*	
("ethnic minority student*" OR («ethnic background» AND student*) OR "ethnic minority pupil*" OR «ethnic background» AND pupil*)) AND («first language» OR L1 OR «mother tongue» OR «native language»)	72 results 8 results assessed verbally 1 abstract viewed 0 full texts viewed

*Note: the advanced query constructed for P4 is missing a left parenthesis here: OR **(**«ethnic background» AND pupil*). Tests done after the data collection revealed that the query returned the same number of results both with and without the parenthesis.

8 results were assessed verbally during the AQS, of which 1 abstract and 0 full text document were viewed. The abstract belonged to a result that was unique to the AQS.

5.2.5 Participant 5 (P5)

P5 was working on a master thesis concerning the concept of oral communication, and how this is presented in Norwegian language textbooks. P5 felt confident in knowledge on this phenomenon, having completed two previous assignments on the subject. The thesis supervisor had provided P5 with several pages of references considered relevant for the project. At the time of the pre-search interview, P5 was unsure of what other sources might be needed but was still looking for any literature on the topic. Research articles or master theses were both acceptable as source types. P5 had not done any searches on the topic recently and had not used the ERIC database prior to data collection. The current information need and intended use of found information for P5 were interpreted in this way:

Information need	**Intended use**
Literature describing or examining the concept of oral communication in textbooks	Inform P5's own methodology and theoretical basis for the master thesis

As P5 stated that the topic was well known from previous course work, the information need could be considered as being a variable CIN, but with a leaning towards a muddled (MIN) type since there was a stated uncertainty as to what additional information was needed in the present situation. The draft version of the SQS query was constructed based on the two facets identified by the author and was run as presented below.

P5 Simple Boolean query	
«oral communication» AND textbooks	576 results 2 results assessed verbally 2 abstracts viewed 0 full texts viewed

2 of the retrieved results in the SQS were assessed, of which 2 abstracts and 0 full text documents. The draft version of the advanced query included 'speech skill*' as a related term to 'oral communication' and truncated versions of 'textbooks' and 'curriculum'. 'Curriculum' was accepted as closely enough related to the present information need by P5. The participant added 'listening skill*' as a term to the AQS query, as this was a commonly overlooked, but important aspect of oral communication as considered by P5.

P5 Advanced Boolean query	
(«oral communication» OR «Speech skill*» OR «listening skill*») AND (textbook* OR curricul*)	5711 results 4 results assessed verbally 4 abstracts viewed 0 full texts viewed

4 results were assessed verbally during the AQS, of which 4 abstracts and 0 full text documents. 1 abstract belonged to a result that was assessed in the SQS, where the abstract also had been viewed. 1 abstract belonged to a result that was not assessed in the SQS, and 2 where unique to the AQS.

5.2.6 Participant 6 (P6)

P6 wanted to explore how women are represented in textbooks and curricula and stated that research-based information on the subject was very welcome as it was not that well known to P6 from before. P6 had searched Oria for sources on the topic recently and had not used the ERIC database prior to data collection. The current information need and intended use of found information for P6 were interpreted as the following:

Information need	Intended use
Any sources on representation of women in literature, including research on women representation in textbooks	To gain oversight and knowledge on chosen research theme

P6's information need may be characterized as a variable MIN, as the topic was relatively new to the participant, and the intention was to gain oversight. A choice was made to construct the SQS query with the term 'textbooks' rather than 'literature' representing one of the facets, as this was considered more precise as to the stated theme for P6's thesis. P6 confirmed that the terms used were familiar, and that their combination seemed appropriate given the present information need.

P6 Simple Boolean query*	
"gender representation" AND textbooks	18 results 10 results assessed verbally 4 abstracts viewed 0 full texts viewed

* The SQS query was limited to 'textbooks', but in hindsight, it should ideally have included 'literature' given the stated information need.

10 of the retrieved results in the SQS were assessed, of which 4 abstracts and 0 full text documents were viewed. The draft version of the advanced query included related terms to 'gender representation', and 'curriculum' as a related term to textbooks. 'Curriculum' was accepted as closely enough related to P6's present information need by the participant.

P6 Advanced Boolean query	
("gender representation*" OR "women representation*" OR "female representation*") AND (textbook* OR curricul*)	38 results 9 results assessed verbally 2 abstracts viewed 0 full texts viewed

9 results were assessed verbally during the AQS, of which 2 abstracts and 0 full text documents. One abstract belonged to a result present but not assessed in the SQS, the other was unique to the AQS.

5.2.7 Participant 7 (P7)

The topic for P7's master thesis was the use of morning circle groups involving the entire class, at the first levels of primary education. P7 was motivated by experiences gained through earlier work and had knowledge on the topic from these encounters. P7 was currently looking for research on the topic and related sub-topics, both for the theoretical foundation for the thesis, and concerning methodology. P7's thesis supervisor had recently provided a few relevant references**.** P7 had not done recent information searches before the data collection and had only used the ERIC database once, during the university library workshop mentioned in section 4.2. The information need and intended use was interpreted in the following manner:

Information need	**Intended use**
Research on the topic of morning circle groups	Inform theoretical foundation and methodology in the thesis

The need was considered as a variable CIN, the topic being stated as familiar by the participant. In mapping out term to be used in the queries for P7, it became clear that both "morning circle groups" or "circle groups" returned few results in ERIC, and no results when combined with "primary school". The term "classroom community" was considered by the author to apply to the goal implicit in using such groups i a school setting and was therefore used instead. The query was presented and explained to P7 via shared screen directly after the pre-search interview. P7 confirmed that the terms used were familiar, and that their combination seemed appropriate given the present information need.

P7 Simple Boolean query	
"classroom community" AND "primary school"	7 results 5 results assessed verbally 1 abstract viewed

	0 full texts viewed

5 of the retrieved results in the SQS were assessed, of which 1 abstract and 0 full text documents were viewed. The draft version of the advanced query included a truncation of the 'classroom community' term, and synonyms and related terms to 'primary school'. No changes were made.

P7 Advanced Boolean query	
"classroom communit*" AND ("primary education" OR "elementary education" OR "primary school*" OR "elementary school*")	230 results 11 results assessed verbally 1 abstract viewed 0 full texts viewed

11 results were assessed verbally, of which 1 abstract that was unique to the AQS.

To sum up, of the 149 results that were assessed, a total of 42 abstracts with 5 related full texts were viewed. 6 of the 7 participants had not used the ERIC database through EBSCO*host* before contributing to the data collection. No changes were made to the proposed SQS queries. In two of the proposed AQS queries a related term was added by the participant, for one of the major facets in each query. See Appendix F for a tabular overview of participants, information needs and queries.

As seen in the presentations above, most of the initial needs were considered as variable CIN types. The interpretation was based on participants stating that the chosen topic was familiar, and thereby represented "known subject matter or domain" in the words of Ingwersen (2000, p. 164). One need could be considered being slightly more muddled (P5), and one variable muddled (P6) owing to the topic being stated as not that familiar.

5.3 Verbalizations of relevance criteria

As described in section 4.5, and shown in the procedures in Appendices D and E, verbalizations of relevance criteria were collected through think-aloud interviews with each participant. To start the think-aloud sessions, variations of the following initial question were

asked by the author after the result lists were retrieved in each SQS and AQS: “How would you judge the relevance of these results? Could you please give an assessment of the results you are seeing on the screen right now?”. This typically triggered responses that concerned immediate impressions with the top results in each list. Care was taken not to co-interpret results. Rather than commenting the result itself, efforts were made to comment and interact on the participant’s utterances: “Why was that reference relevant/not relevant?” And confirming participant’s assessment: “So, it’s the choice of methods in this paper that are the most interesting to you?” and similar. All participants assessed results within the top three of the lists in the SQS. In the AQS, this only applied to four of the participants as assessments were distributed more disparately across rankings.

Depending on the responses to the initial question, the author (A) would follow up with general comments such as “Can you tell me a bit more about why this result is relevant to you?”, or remarks adjusted to the response from the participant. One such example of adjustment is shown in this segment from P1’s SQS, during assessment of the title ‘Playification of the Curriculum: Learnings from Collaborative Classroom Research’:

> P1: “That one, number 22, that resonates well with me”
>
> A: “Yes, is that a term that you know well from before? Playification?”
>
> P1: “Yeah, so what I’m picturing here is, when I’m reading that title, is that it involves a play based approach to the curriculum? So, I’m thinking that this would absolutely be useful to have”

With regards to coding, the last verbalization by P1 in this segment was coded to the category ‘Value in use’ (Use or situational match – Human characteristic) and categories ‘Topic’ (Content – Object characteristics) and ‘Title’ (Object – Object characteristics). Adjusted interventions by the author such the one above would generate explanations from participants, although quite often only a minimal, affirmative confirmation such as “Yes?” or “Mhm” would trigger detailed comments. One such example is this segment from P5’s SQS, where an abstract is being examined:

> P5: “Uhm, something which is nice, although not necessarily because it is *that* relevant to me, is that they [the authors of the assessed article] are looking at how they are working with feedback”

A: “Mm”

P5: «Yeah, because I’m not looking at that, but the fact that they are looking at how these feedback comments are given in textbooks”

A: “Mm”

P5: “That makes me think that I can look through their article and find theory in that article, which I then can use in my master’s thesis”

The first verbalization was coded to ‘Appropriateness to task’ (Use or situational match – Human characteristic) and ‘Abstract’ (Object – Object characteristics), the second to ‘Methodology’ (Content – Object characteristics) and ‘Abstract’ (Object – Object characteristics), and the third to ‘Value in use’ (Use or situational match – Human characteristic) and ‘Abstract’ (Object – Object characteristics).

The purpose of these examples is to show how frequency of criteria in verbalizations were derived. The two examples of adjusted intervention and minimal affirmative confirmations above, are considered as representing two ends of a spectrum in terms of interventions on the authors part during think-aloud interviews. As anchor examples for additional coded verbalizations are provided previously in section 4.10.3, further examples are not included here, in consideration of space.

5.4 Frequency of stated relevance criteria

The total number of verbalizations coded to each criterion was derived via the Matrix Coding Query function in NVivo described in section 4.10.3. As a reminder, the object characteristics group contained a total of 30 single criteria, and the human characteristics group 12. The reader is referred to section 4.10.3, as well as the codebook in Appendix G for a description of each criterion.

5.4.1 Criteria stated in the SQS

The frequency of verbalized criteria in the simple query session are shown below in Table 11**:**

Simple Query Session				
Criteria groups	**Total number of criteria coded in verbalizations**	**Percent**	**Participants**	**Percent of participants**
Object characteristics	**238**	**72,8 %**		
Content	112	34,3 %	7	100 %
Object	118	36,1 %	7	100 %
Validity	8	2,5 %	3	42,8 %
Human characteristics	**89**	**27,2 %**		
Use or situational match	37	11,3 %	7	100 %
Cognitive match	29	8,9 %	6	85,7 %
Affective match	22	6,7 %	5	71,4 %
Belief match	1	0,3 %	1	14,2 %
TOTAL	**327**	**100 %**		

Table 11: Relevance criteria stated by participants in the SQS. Hybrid based on Barry (1994, p. 157) and Saracevic (2017, pp. 57-58).

Criteria within all of the generalized groups proposed by Saracevic (2017, pp. 57-58) were considered present in the SQS, although only one such instance was coded to the Belief match-group. Likewise, verbalizations considered to be containing criteria in the 'validity' group were relatively few (8). To add some nuance to these numbers, the four most frequently coded object and content criteria in the SQS are displayed in the table below:

Simple Query Session		
Object characteristics groups	**Total number of criteria coded in verbalizations**	**Percent of criteria within category**
Content	112	
Treatment	23	20,5 %
Geographic location of study	20	17,9 %
Currency (publication date or year)	19	17,0 %
Topic	16	14,3 %
Object	118	
Title	55	47,0 %
Abstract	22	18,6 %
Full text	18	15,3 %
Subject terms	8	6,8 %

Table 12: Most frequent object criteria in the SQS

The most frequently coded content criteria, 'Treatment' is understood as closely connected to the fourth most frequent, which is 'Topic'. If treated as a subcategory in that it is defined as they way in which the topic or subject of a document is being discussed, these two topic-related criteria account for over a third of those in the content group. Geographic location and currency accounted for approximately a fifth of the content criteria each. Title is by far the most used criterion in the object group, followed by abstract and full text. The reason that the number of coded instances of the title criterion (55) does not directly reflect the number of retrieved results in the SQS (48), is that title was used as a criterion for assessment more than once for some of the results.

The table below shows the frequency of criteria coded in the human characteristics groups, excluding 'Belief match' which contained one instance:

Simple Query Session		
Human characteristics groups	**Total number of criteria coded in verbalizations**	**Percent of criteria within category**
Affective match	22	
Uncertainty	14	63,6 %
Curiosity	6	27,3 %
Enthusiasm	2	9,0 %
Cognitive match	29	
Understanding	14	48,3 %
Link to previous knowledge	14	48,3 %
Novelty	1	3,4 %
Use or situational match	37	
Appropriateness to situation, or task	18	48,6 %
Value in use	18	48,6 %
Urgency	1	2,7 %

Table 13: Most frequent human criteria in the SQS

Uncertainty dominated as the emotional response in the SQS, followed by curiosity. These can however be considered as closely associated, as will be discussed in subsequently in

section 5.5.1. In the cognitive match group, understanding and link to previous knowledge were the most prominent criteria. These two can also be seen as closely linked, as understanding necessitates previous knowledge. The latter code is however used to differentiate such verbalizations when a certain example of this knowledge, such as a particular theory or concept, is mentioned during the assessment. 'Appropriateness to task' and 'Value in use' were each coded 18 times in the situational match category.

5.4.2 Criteria stated in the AQS

Table 14 below shows the frequency of verbalized criteria in the advanced query session:

Advanced Query Session				
Criteria groups	**Total number of criteria coded in verbalizations**	**Percent of verbalizations**	**Participants**	**Percent of participants**
Object characteristics	240	73,6 %		
content	130	39,9 %	7	100 %
object	104	31,9 %	7	100 %
validity	6	1,8 %	3	42,8 %
Human characteristics	86	26,4 %		
use or situational match	31	9,5 %	6	85,7 %
cognitive match	12	3,7 %	5	71,4 %
affective match	40	12,3 %	6	85,7 %
belief match	3	0,9 %	2	28,5 %
TOTAL	326	100 %		

Table 14: Relevance criteria stated by participants in the AQS. Hybrid based on Barry (1994) and (Saracevic, 2017, pp. 57-58).

All criteria groups were also deemed present in the transcripts from the AQS when viewed as a whole, and the distribution of human and object characteristics closely resembled that of the SQS. However, one difference is that none of the human characteristics groups were considered as being mentioned by all seven participants during the AQS. Table 15 gives an overview over most frequently mentioned object criteria:

Advanced Query Session		
Object characteristics groups	**Total number of criteria coded in verbalizations**	**Percent of criteria within category**
Content	130	
Treatment	46	35,4 %
Geographic location of study	24	18,5 %
Topic	21	16,2 %
Currency (publication date or year)	13	10 %
Object	103	
Title	56	53,4 %
Abstract	18	17,5 %
Title recognized from simple query	14	13,6 %
Subject terms	8	7,8 %

Table 15: Most frequent object criteria in the AQS

Treatment and topic are the dominant criteria also in the AQS. When seen as combined these two account for 51,2%, followed by geographic location of study, and currency. Content criteria were mentioned slightly more often in the AQS than the SQS.

Advanced Query Session		
Human characteristics groups	**Total number of criteria coded in verbalizations**	**Percent of criteria within category**
Affective match	40	
Uncertainty	19	47,5 %
Curiosity	15	37,5 %
Enthusiasm	3	7,5 %
Cognitive match	12	
Understanding	6	50,0 %
Link to previous knowledge	4	10,0 %
Novelty	2	5,0 %
Use or situational match	31	
Appropriateness to situation, or task	24	77,4 %
Value in use	6	19,4 %
Urgency	1	3,2 %

Table 16: Most frequent human criteria in the AQS

'Uncertainty' was also the dominant affective match criteria in the AQS, followed by 'Curiosity'. A large increase occurred in the frequency of affective match criteria between searches, which almost doubled from 22 in the SQS to 40 verbalizations in the AQS. This is in large part due to the number of instances where 'Curiosity' was coded, which increased from 6 to 15. Cognitive match verbalizations decreased from 29 in the SQS to 12 in the AQS, with almost equal decrease in 'Understanding' and 'Link to previous knowledge'.

In Table 17, total number of verbalized criteria are shown, which is the basis for the analysis in the next section.

Both Query Sessions				
Criteria categories	**Total number of criteria coded in verbalizations**	**Percent of verbalizations**	**Participants**	**Percent of participants**
Object characteristics	**478**	**73,2 %**		
content	242	**37,0 %**	7	100 %
object	222	**33,9 %**	7	100 %
validity	14	**2,1 %**	4	57,1 %
Human characteristics	**175**	**26,9 %**		
use or situational match	68	**10,4 %**	7	100 %
cognitive match	41	**6,3 %**	7	100 %
affective match	62	**9,5 %**	6	85,7 %
belief match	4	**0,6 %**	2	28,5 %
TOTAL	**653**	**100 %**		

Table 17: Total number of criteria stated by participants. Hybrid based on Barry (1994, p. 157) and Saracevic (2017, pp. 57-58).

5.5 Frequency of co-occurring relevance criteria

Table 12 shows the result of another matrix coding query, which was run to establish what relevance criteria that were used to code the same verbalizations in the think-aloud transcripts. This was done to identify what criteria in the 'Human' and 'Object' main groups that were identified and coded most often and how they intersected in verbalizations. Frequencies of coding could subsequently indicate relationships between criteria, and how these were distributed across categories. The heatmap in Table 18 shows the categories of

criteria that were interpreted as mentioned concurrently by participants and frequency of these co-occurrences.

Human characteristics	Object characteristics		
	Content	Object	Validity
Affective match	11	31	0
Curiosity	3	6	0
Empathy	0	1	0
Enthusiasm	1	3	0
Fun	0	1	0
Uncertainty	7	20	0
Belief match	3	1	1
Personal credence given to information	3	1	1
Cognitive match	10	19	2
Link to previous knowledge	5	7	2
Novelty	2	1	0
Understanding	3	11	0
Use or situational match	27	13	2
Appropriateness to situation, or task	17	9	1
Urgency	2	1	0
Value in use	8	5	1

Table 18: Distribution of intersecting codes in the think-aloud transcripts (no. of verbalizations with intersecting codes)

To give an overall impression of the presence of co-occurring relevance criteria, the table above only includes combined counts for the object groups. A closer analysis of the most frequent co-occurrences between human and object characteristics groups and what criteria are involved in these interactions is given below. These interactions and interdependencies of categories and groups can give insight with regards to manifestations of different relevance types.

5.5.1 Affective match and co-occurring object characteristics

As can be seen in Table 12, the most frequently coded 'Affective match' criteria was 'Uncertainty'. This co-occurred most often with the 'Object' criteria 'Title' (15 of 20 co-occurrences). Examples of these interactions often involved participants questioning the meaning of certain terms in titles:

> "Hm. This could be relevant, but then again not. Because, uhm, it concerns this concept of 'self regulated learning' in a digital age" (P2, AQS)

> "What do they mean with 'language planning'?" (P6 AQS)

Verbalizations that indicated uncertainty were sometimes considered as closely linked to the category 'Curiosity', and therefore challenging to categorize. Some verbalizations were coded to both of these criteria, in instances where uncertainty regarding the meaning of a term was connected to a stated action intended to relieve doubt:

> «I would have accessed that one just to check what 'ELT' meant, to make sure I did not miss anything..." (P6 AQS).

The most frequently co-occurring 'Content' category with 'Uncertainty' was 'Treatment'. One example of such interactions from P7's AQS:

> "Is this regarding students in... is that a certain type of school? 'Caring school'?"

Uncertainties like these would sometimes be resolved when participants continued the assessment of a result, or they would remain as the participant would decide to go on to the view additional results.

5.5.2 Belief match and co-occurring object characteristics

The only identified criteria to the group 'Belief match' was 'Personal credence given to information', which co-occurred twice with the 'Content' criteria 'Geographic location of study'**,** discerned from the 'Object' criteria 'Subject terms'. in P1's AQS:

> "And then when I see 'New Zealand', I know that we have gotten a lot from there"
>
> "And just because it says, 'New Zealand' here, right, then I'm thinking 'Oh, I want to read this'".

P1 explained that research on the topic of play-based learning originating in this country was often adopted in a Norwegian context.

5.5.3 Cognitive match and co-occurring object characteristics

'Understanding' was coded most often within this group, and it co-occurred most frequently with the 'Object' criteria 'Title' (6 of 11). One such example from P5's SQS:

> «This one, without having read it, I see straight away that it says 'e-textbook versus paper textbook', so I'm thinking that it would [...] It would be nice to have a look at that, considering that chapter"

The chapter mentioned at the end of the quote is a chapter planned as part of P5's master thesis. This mention of the participants work task at the end of the quote indicates a transition from a cognitive match of understanding (in the interpretation of the title of the result), into a situational match in terms of potential usefulness.

As previously mentioned, understanding is considered as closely connected with link to previous knowledge. The following verbalization is an example of this criterion co-occurring with title, from P4's AQS:

> "'Hegemony', that's also a term which we [the students] have learned [...]"

This is one example of a term that was connected to themes discussed as part of a university course.

5.5.4 Situational match and co-occurring object characteristics

'Appropriateness to task' was the most frequently coded situational match-criterion, co-occurring most often with the content criterion 'Treatment'. Below is an example from P1's AQS, where the assessment is made based on the abstract of a result:

> "Yeah, ok, this one is very pertinent, because it deals with in-depth learning [...] through play?"

The use of the word 'pertinent' in the translation from Norwegian ('aktuell'), refers to the participant underlining the situational aspect, i.e., appropriateness. 'Value in use' co-occurred most often (three times) with 'Methodology', here derived from the assessment of an abstract:

> «They way this, this paper handles this, is in many ways an approach I also would consider" (P4 AQS)

A connection can be seen as made here, to the ongoing work task and choice of methodology in the participant's upcoming project.

5.6 Differences between the two queries

Most of the attention in the results section so far has been devoted to the frequencies of relevance criteria stated during think-aloud interviews. To supplement this, a short summary of participants' immediate thoughts and reflections regarding the differences between the two queries (Q9 in the interview guide) is included here. Responses were varied: The

advanced query was considered useful by some of the participants, in that it was considered to improve the number of relevant results (P1, P4, P7) and increased the chances of encountering new perspectives on the topic (P2, P4, P7). Others preferred the simple version, in that it was perceived as more precise and to the point, reducing the amount of potentially irrelevant results (P5, P6), and because the participant felt that enough information was found for the time being, after the first search concluded (P6). The two queries were also considered as quite similar in usefulness by some (P3, P4).

6 Discussion

The background for this project relies on the idea that Boolean operators might not be perceived as useful by graduate students. In other words, that they might be quite satisfied with the results of searches where such operators are not applied, and consequently do not feel the need to learn using them. Research reviewed earlier in section 3.1 support the assumption that students find Boolean operators challenging to use. The findings in the two papers by Lowe et al. (2018; 2020) also show that the difference in performance of natural language versus simple Boolean, and simple versus advanced Boolean, might not be that much when measured as topical relevance of results.

At the outset of this project, the ambition was to examine this idea through an explicit comparison of the difference in subjective relevance of the results, from a user-generated query and an expanded, Boolean query. As shown in the description of the pilot test in section 4.6, a fundamental challenge to this ambition was that of creating a research design and data collection procedure that enabled the observation of user-generated queries and search activity. In the revised research design leading to the results presented in the previous sections, focus is instead directed at the participants' relevance assessments of results from a simple and an advanced version of a Boolean query, constructed on the basis of each participant's information need.

The concept of perceived usefulness pertains to the situational context of the user – and whether the retrieval technique represented by Boolean queries helps with the accomplishment of a work task. In this project, the work task is represented by each participant's master thesis project, a task which is the construction stage for all participants.

Further, this work task is understood as the context which gives rise user's ASK and subsequent information needs, as proposed by Ingwersen (1996, p. 15). Insight into what properties that characterize each participant's information need was attempted through the inclusion of context, expressed as participants' familiarity with topic and intended use. The typology from Ingwersen (2000) and the connection to the task concept presented in section 2.4/2.4.1 was then used to categorize each participant's need.

6.1 Perceived usefulness of the two Boolean queries

When looking at the immediate responses gathered in the post-search interview, 4 of the 5 participants that were considered as starting the search session with a variable CIN type need (P1, P2, P4, P7), commented that the advanced query was valuable, in that it provided an increased number of relevant results, and made it possible to explore new perspectives on the topic in question. The two participants that were considered to start the sessions with more muddled type needs (P5, P6) stated that the simple version was preferrable, due to precision and anticipated reduction in potential irrelevant results. A connection might be seen in these result, to the findings from Vakkari et al. (2003), where increased subject knowledge was related to increased numbers of assessed as partially relevant. These results are only based on summaries of reactions collected immediately after the searches were finished, representing a small part but important of the full picture.

When only this small part of the picture is provided, a somewhat straightforward answer to RQ1 would be that graduate students perceive the usefulness of Boolean queries according to the degree of how well defined and variable the present information need is at any given time. If the knowledge domain is well known (as in a variable CIN need), a student is assumed to be relatively comfortable with assessing retrieved results in the light of existing knowledge, and therefore welcoming novel or unexpected perspectives even though relevance assessments can be influenced by uncertainty. It can then be argued that the use of synonyms and related terms in an expanded query accommodates the intention of a conscious topical need, as such queries aims at widening the scope with regards to what concepts and perspectives that are used to describe the topic in question. This echoes the potential significance of partially relevant results in the early stages of a search, found by Spink et al. (1998).

To add nuance to this understanding, the compromise in Taylor's (1968) Q4 need must be added to the picture. It is argued that the usefulness of a query to some degree also will be perceived in light of the effort needed to reach this compromise. Even though a variable CIN type need signifies a well-known subject, there is of course still a matter of identifying what terms are the most useful or appropriate to include in a query, and to do so with the correct use of Boolean operators. The findings from Pennanen and Vakkari (2003) provide an empirical basis for the claim that increased subject knowledge heightens the user's ability to locate useful terms. If the work task demands comprehensives in a search, the effort needed to locate and combine terms is arguably larger, as more terms are needed. Requirements for comprehensiveness are seen in some of the needs, when the goal is stated as locating research gaps as a means of positioning the thesis project (P1, P2, P4).

To further contextualize the information needs of the participants and incorporate RQ2, the criteria verbalized in the think-aloud interviews are drawn into discussion. As shown, 'Uncertainty' was the most frequently occuring affective match criteria in both search sessions, and it arose in many cases because of doubt as to what certain terms really meant. This may be seen as connected to uncertainty in topicality and cognitive relevance assessments associated with the variable CIN needs as described in Ingwersen (2000, p. 165). The language aspect seems to play a role in this uncertainty, as it becomes unclear whether the encountered concepts are in fact new or not. Typical 'problematic' concepts observed in in this setting of the educational field of research were 'curriculum' and 'K12', as these were challenging for participants to map to a synonym in Norwegian.

6.2 Manifestations of relevance

The results presented in section 5.5 provides snapshots of what stated criteria the participants relied on during one IR session, and how these were interpreted by the author through the deductive coding process and subsequent analysis. Looking at these results, this part of the discussion concerns relations between attributes of relevance and relevance manifestations proposed in the overview by Ingwersen (2000, p. 13) shown in Figure 5, page 28, and Borlund's model (2003, p. 915), repeated here for easier reference.

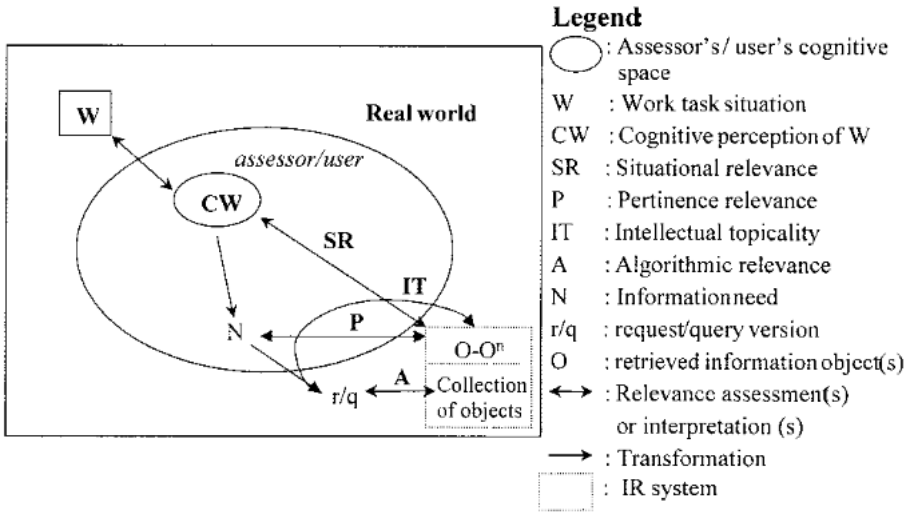


Figure 11: Main types of relevance relationships involved in an IR situation, from Borlund (2003, p. 915) (repeated)

The inference attribute of Topical relevance (intellectual topicality) is argued to be exhibited in the results, in the cases where topic is inferred at a semantic level. For example, when uncertainty arises as a result of reading terms of a title that do not immediately give meaning. Intellectual topicality then connects further to pertinence/cognitive relevance, as the uncertainty becomes related to whether or not the result is relevant to the information need or intentions. The affective relevance type is accordingly present in both instances, as illustrated by the top row in the tabular overview.

Situational relevance assessments are present in the results, when the work task situations are directly drawn in to support the assessment (intention and relation attributes in the tabular overview), and the participant focuses on the potential usefulness (inference attribute), e.g., in deciding on the choice of methodology. Assessments of situational relevance can be connected to changes in information needs, as argued by Borlund (2003, pp. 922-923). In this context, situational relevance is considered key in understanding the perceived usefulness of a certain retrieval technique.

6.3 Suggestions for pedagogical practice

Based on the findings and discussion in this thesis, the interaction between information need type and efforts needed to compromise this need, can be seen as a fundamental

background in understanding students' perceived usefulness of Boolean queries. In instruction and pedagogical approaches, a recognition of these aspects will be beneficial.

Interactions between different groups of relevance types shown in section 6.1 are complex but provide insight into how assessment processes in the early stages of a user's project might be understood from a librarian/intermediate point of view. One suggestion could be to heighten the awareness for both users and intermediaries, as to what criteria are being applied when a result is retrieved.

6.4 Future context for the role of Boolean search strategies

Saracevic (1997) stated that research and development efforts often either did not properly acknowledge or did not recognize that IR systems were actually designed to meet human needs, and that interaction with all its complexity and messiness was not being dealt with. A quarter of a century later, the impression now might be that interaction is indeed being investigated and dealt with to a large degree but that in many cases, the way it is being done and the consequences for the end user, are unknown to the public. The accumulation of big data on user interactions on social media with resulting algorithmic response is one such example. Another is the application of artificial intelligence (AI) and machine learning in modeling the end-user, for example by the automated recommendations for references during the writing process by services such as Keenious[9].

The number of recommendation-based services in the user's search environment is also increasing, with built-in functions such as the bX article recommender in Oria, and in Clarivate's new interface for the Web of Science database launched in 2021. What is the future of Boolean search strategies in this context? It is argued that the level of potential precision offered by Boolean search is powerful and can provide students with control, given enough training and willingness to employ these techniques. The findings in this thesis show that the efforts needed to construct these queries should be seen in the context of user's need, and examined through situational relevance of results, not only topical relevance.

6.5 Limitations

Lengthy considerations preceded the choice to carry on with a user-centered study in the context of an unpredictable pandemic situation. As all lecture and laboratory activity on

[9] https://keenious.com/

campus was canceled for the main part of 2020, the potential participants in this study had already endured almost two full semesters in an all-digital study situation at the time of recruitment. In this respect, the digital conversational situation had established itself as an activity of the everyday, in great contrast to how these situations might have been experienced a year earlier before the pandemic. At the same time, the concept of 'Zoom fatigue' had also established itself in mainstream media discourse and in the research literature, see e.g., Bailenson (2021), signifying the more negative aspects of this situation.

Still, the prospect of collecting data for a user-centered IIR-study completely via digital means also posed an interesting challenge within the frame of a master thesis project, which in turn gave the author (and hopefully, the reader) interesting and valuable insight on how video conference software can be utilized for these purposes.

Another source of much deliberation was the construction and quality of the Boolean queries used in the search sessions. They represent an effort on the authors part to reflect the dynamic information needs of the participants while at the same time allowing for the testing of both a 'simple' and 'advanced' approach to constructing such queries. Readers will perhaps react or even object to the number of exact phrases used in the advanced version, or the choice of synonyms and related terms. There is of course also the risk that the queries poorly matched the real information needs of the participants. Queries were constructed for the purpose of reflecting each participant's information need, and for testing. It can be discussed whether they come across as realistic, and whether participants found them acceptable because they were already prepared, and because of the setting. The fact that two of the participants chose to include terms at their own will mitigates this to a certain extent.

Another interesting and challenging aspect of this project was the deductive coding procedure, which sought to identify what criteria the participants relied on. This was not a straightforward procedure at all, as verbalizations contained multiple layers of interpretations and elements that were often difficult to categorize. A revised process for this part of the data analysis including more than one coder to develop a more robust deductive procedure, would be an exciting future project to undertake.

6.5.1 Mediated search as a data collection method

In relation to choosing mediated searches as a method for data collection this quote was cause for much consideration: “If the interaction is mediated involving an intermediary, still another complex stratum, is added, very interesting in itself.” (Saracevic, 1996a, p. 6). This complex stratum, as Saracevic puts it, is certainly present within the research design of this thesis, both on the user side (the participant is interacting with the author both in the negotiation of the information need formulated through a query, and through think-aloud statements during relevance assessment) *and* on the system side (the surface level may be considered as doubled by the Zoom video meeting service software).

During data collection, the author forms part of the participants’ information seeking context, becoming an (asymmetric) conversational partner during interpretations of search results, and therefore possibly a mitigating factor in terms of uncertainty in the participant. This might have influenced the validity of results with regards to bias. The Zoom software was as argued above, relatively familiar to the participants in question. However, the remote-control function and the ERIC database was not, possibly affecting the data collected.

7 Conclusions

The combination of increasingly ubiquitous, user friendly and familiar discovery services and recommendation systems, and the observed similarity of topical relevance in results from both simple and advanced Boolean queries in Lowe et al. (2020) can be used as arguments for a shift in pedagogical approaches and a change of focus in library instruction away from Boolean search. Such a change of focus implies considerations with regards to what students (level and area of study) need to learn advanced Boolean search techniques. The present master thesis sought to add depth to such considerations, by incorporating the students’ perspective**.**

In the reported project, this perspective is represented through the information needs belonging to the participants, and their subsequent relevance assessments. The findings suggest that advanced Boolean queries can be useful in subject database searching if the information need is rooted in a well-known subject domain for the user, but that a compromise is always present when the need is formulated as a query.

7.1 Future research

The empirical data underlying this thesis is limited to a single search session per participant. Developing this project into a longitudinal study would provide more insight into how relevance assessments, relevance criteria, and inherently also query preference develops over time. This would be especially interesting if coordinated to run parallel to participants' timeline for an entire master thesis project or similar. Such a project could potentially also make use of measures beyond that of reported relevance and criteria in a result list by incorporating actual use of results in the final exam text or similar, as in the studies by Vakkari et al. (2019). In terms of data collection procedures for a longitudinal project, a less obtrusive and time-consuming approach than think-aloud interviews could be favorable.

An interesting approach for future research on this theme could also be a user-centered comparative study looking at the performance of recommendation-based systems such as the Keenious service, and a more traditional query-based approach. Perceived usefulness simple and advanced Boolean queries measured through situational relevance warrants further investigation, ideally in a longitudinal study of user groups.

Literature

Assarroudi, A., Heshmati Nabavi, F., Armat, M. R., Ebadi, A., & Vaismoradi, M. (2018). Directed qualitative content analysis: the description and elaboration of its underpinning methods and data analysis process. *Journal of Research in Nursing*, *23*(1), 42-55. https://doi.org/10.1177/1744987117741667

Bailenson, J. N. (2021). Nonverbal overload: A theoretical argument for the causes of Zoom fatigue. *Technology, Mind, and Behavior*, *2*(1). https://doi.org/10.1037/tmb0000030

Barry, C. L. (1994). User-defined relevance criteria: An exploratory study. *Journal of the American Society for Information Science*, *45*(3), 149-159. https://doi.org/10.1002/(SICI)1097-4571(199404)45:3<149::AID-ASI5>3.0.CO;2-J

Belkin, N. J. (2008). Some(what) grand challenges for information retrieval. *SIGIR forum*, *42*(1), 47-54. https://doi.org/10.1145/1394251.1394261

Belkin, N. J., Oddy, R. N., & Brooks, H. M. (1982a). ASK for information retrieval: Part I. Background and theory. *Journal of Documentation*, *38*(2). http://infobehavior.pbworks.com/f/Belkin%20ASK%20p1.pdf

Belkin, N. J., Oddy, R. N., & Brooks, H. M. (1982b). ASK for information retrieval: Part II. Results of a design study. *Journal of Documentation*, *38*(3). https://surface.syr.edu/cgi/viewcontent.cgi?article=1151&context=istpub

Borlund, P. (2003). The concept of relevance in IR. *Journal of the American Society for Information Science and Technology*, *54*(10), 913-925. https://doi.org/10.1002/asi.10286

Borlund, P. (2010). The Cognitive viewpoint: The Essence of Information Retrieval Interaction. In B. Larsen, J. W. Schneider, F. Åström, & B. Schlemmer (Eds.), *The Janus-faced scholar: A festschrift in honour of Peter Ingwersen* (pp. 23-34). Det Informationsvidenskabelige Akademi. https://www.lunduniversity.lu.se/lup/publication/c4791b4c-82d1-4dab-8d5a-363e11f8d8d6

Borlund, P., & Dreier, S. (2014). An investigation of the search behaviour associated with Ingwersen's three types of information needs. *Information Processing & Management*, *50*(4), 493-507. https://doi.org/10.1016/j.ipm.2014.03.001

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Byström, K., & Hansen, P. (2005). Conceptual framework for tasks in information studies. *Journal of the American Society for Information Science and Technology*, *56*(10), 1050-1061. https://doi.org/10.1002/asi.20197

Byström, K., & Kumpulainen, S. (2020). Vertical and horizontal relationships amongst task-based information needs. *Information Processing & Management*, *57*(2), 1-14. https://doi.org/10.1016/j.ipm.2019.102065

Catalano, A. (2013). Patterns of graduate students' information seeking behavior: a meta-synthesis of the literature. *Journal of Documentation*, *69*(2), 243-274. https://doi.org/10.1108/00220411311300066

Clough, P., & Sanderson, M. (2013). Evaluating the performance of information retrieval systems using test collections. *Information research*, *18*(2), Article 582. http://informationr.net/ir/18-2/paper582.html#.YaSbmNDMJaQ

Cmor, D., & Li, X. (2012). Beyond boolean, towards thinking: discovery systems and information literacy. *Library Management*, *33*(8/9), 450-457. https://doi.org/10.1108/01435121211279812

Cosijn, E. (2006). Relevance judgements within the context of work tasks. In *Proceedings of the 1st international conference on Information interaction in context* (pp. 20–29). Association for Computing Machinery. https://doi.org/10.1145/1164820.1164827

Cosijn, E., & Ingwersen, P. (2000). Dimensions of relevance. *Information Processing & Management*, *36*(4), 533-550. https://doi.org/10.1016/S0306-4573(99)00072-2

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319-340. https://doi.org/10.2307/249008

Dempsey, M., & Valenti, A. M. (2016). Student Use of Keywords and Limiters in Web-scale Discovery Searching. *The Journal of Academic Librarianship*, *42*(3), 200-206. https://doi.org/10.1016/j.acalib.2016.03.002

EBSCO Connect. (2018, October 24). *How can I improve my search results?* https://connect.ebsco.com/s/article/How-can-I-improve-my-search-results?language=en_US

EBSCO Connect. (2019, March 16). *What is the Apply Equivalent Subjects expander?* https://connect.ebsco.com/s/article/What-is-the-Apply-Equivalent-Subjects-expander?language=en_US

EBSCO Connect. (2020, January 13). *How is relevance ranking determined in EBSCOhost?* https://connect.ebsco.com/s/article/How-is-relevance-ranking-determined-in-EBSCOhost?language=en_US

EBSCO Connect. (2021, August 17). *Applying Search Modes in EBSCOhost and EBSCO Discovery Service*. https://connect.ebsco.com/s/article/Applying-Search-Modes-in-EBSCOhost-and-EBSCO-Discovery-Service?language=en_US

EBSCOhost. (n.d.-a). *About the database: ERIC*. Retrieved 27.10.2020 from http://support.ebsco.com/help/?int=ehost&lang=en&feature_id=Databases&TOC_ID=Always&SI=0&BU=0&GU=1&PS=0&ver=live&dbs=,eric#A

EBSCOhost. (n.d.-b). *Including Phrases in a Search*. Retrieved 02.07.2021 from https://support.ebsco.com/help/index.php?help_id=46

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: verbal reports as data* (Rev. ed.). MIT Press.

Frické, M. (2021). Boolean Logic. *Knowledge Organization*, *48*(2), 177-191. https://doi.org/10.5771/0943-7444-2021-2-177

Georgas, H. (2013). Google vs. the library: student preferences and perceptions when doing research using Google and a federated search tool. *portal: Libraries and the Academy*, *13*(2), 165-185. https://academicworks.cuny.edu/cgi/viewcontent.cgi?article=1080&context=bc_pubs

Gusenbauer, M., & Haddaway, N. R. (2019). Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research Synthesis Methods*, *11*(2), 181-217. https://doi.org/10.1002/jrsm.1378

Hamlett, A., & Georgas, H. (2019). In the Wake of Discovery: Student Perceptions, Integration, and Instructional Design. *Journal of Web Librarianship*, *13*(3), 230-245. https://doi.org/10.1080/19322909.2019.1598919

Harter, S. P. (1992). Psychological relevance and information science [Article]. *Journal of the American Society for Information Science*, *43*, 602-615. https://doi.org/10.1002/(SICI)1097-4571(199210)43:9<602::AID-ASI3>3.0.CO;2-Q

Hertzum, M., Borlund, P., & Kristoffersen, K. B. (2015). What Do Thinking-Aloud Participants Say? A Comparison of Moderated and Unmoderated Usability Sessions. *International Journal of Human-Computer Interaction*, *31*(9), 557-570. https://doi.org/10.1080/10447318.2015.1065691

Hjørland, B. (2010). The foundation of the concept of relevance. *Journal of the American Society for Information Science and Technology*, *61*(2), 217-237. https://doi.org/10.1002/asi.21261

Hjørland, B. (2015). Classical databases and knowledge organization: A case for boolean retrieval and human decision-making during searches. *Journal of the Association for Information Science and Technology*, *66*(8), 1559-1575. https://doi.org/10.1002/asi.23250

Hsieh, H.-F., & Shannon, S. E. (2005). Three Approaches to Qualitative Content Analysis. *Qualitative Health Research*, *15*(9), 1277-1288. https://doi.org/10.1177/1049732305276687

Høgheim, S. (2020). *Masteroppgaven i GLU [The Master Thesis in Teacher Education]*. Fagbokforlaget.

Ingwersen, P. (1996). Cognitive Perspectives of Information Retrieval Interaction: Elements of a Cognitive IR Theory. *Journal of Documentation*, *52*(1), 3-50. https://doi.org/10.1108/eb026960

Ingwersen, P. (1999). Cognitive information retrieval. *Annual Review of Information Science and Technology*, *34*, 3-52.
Ingwersen, P. (2000). Users in context. In M. Agosti, F. Crestani, & G. Pasi (Eds.), *Lectures on Information Retrieval: Third European Summer-School, ESSIR 2000 Varenna, Italy, September 11-15, 2000. Revised Lectures* (pp. 157-178). Springer. http://peteringwersen.info/publications/4115_users_in_context-italy2000.pdf
Ingwersen, P., & Järvelin, K. (2005). *The Turn: Integration of Information Seeking and Retrieval in Context* (Vol. 18). Springer. https://ebookcentral-proquest-com.ezproxy.oslomet.no/lib/hioa/reader.action?docID=303407
Institute of Education Sciences. (2021). *ERIC Collection Snapshot*. https://eric.ed.gov/pdf/ERIC_Collection_Snapshot.pdf
Jiang, J., He, D., Kelly, D., & Allan, J. (2017). Understanding Ephemeral State of Relevance. In R. Nordlie & N. Pharo (Eds.), *Proceedings of the 2017 Conference on Human Information Interaction and Retrieval* (pp. 137-146). Association for Computing Machinery. https://doi.org/10.1145/3020165.3020176
Kekäläinen, J., & Järvelin, K. (2002). Using graded relevance assessments in IR evaluation. *Journal of the American Society for Information Science and Technology*, *53*(13), 1120-1129. https://doi.org/10.1002/asi.10137
Kelly, D. (2009). Methods for Evaluating Interactive Information Retrieval Systems with Users. *Foundations and Trends® in Information Retrieval*, *3*(1—2), 1–224. https://doi.org/10.1561/1500000012
Kuhlthau, C. C. (2004). *Seeking meaning: a process approach to library and information services* (2nd ed.). Libraries Unltd.
Kuhlthau, C. C., Heinström, J., & Todd, R. J. (2008). The 'information search process' revisited: Is the model still useful. *Information research*, *13*(4), 13-14. http://informationr.net/ir/13-4/paper355.html
Lee, B., & Chung, E. (2016). An Analysis of Web-scale Discovery Services From the Perspective of User's Relevance Judgment. *The Journal of Academic Librarianship*, *42*(5), 529-534. https://doi.org/10.1016/j.acalib.2016.06.016
Lefebvre, C., Glanville, J., Briscoe, S., Littlewood, A., Marshall, C., Metzendorf, M.-I., Noel-Storr, A., Rader, T., Shokraneh, F., Thomas, J., & Wieland, L. S. (2020). 4.4.2 Structure of a search strategy. In J. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. Page, & V. Welch (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions* (Version 6.1 ed.). Cochrane. https://training.cochrane.org/handbook/current/chapter-04#section-4-4-2
Lowe, M. S., Maxson, B. K., Stone, S. M., Miller, W., Snajdr, E., & Hanna, K. (2018). The Boolean is Dead, Long Live the Boolean! Natural Language versus Boolean Searching in Introductory Undergraduate Instruction. *College & Research Libraries*, *79*(4), 517-534. https://doi.org/10.5860/crl.79.4.517
Lowe, M. S., Stone, S. M., Maxson, B. K., Snajdr, E., & Miller, W. (2020). Boolean redux: Performance of advanced versus simple boolean searches and implications for upper-level instruction. *The Journal of Academic Librarianship*, *46*(6), 1-8. https://doi.org/10.1016/j.acalib.2020.102234
Makri, S., Blandford, A., & Cox, A. L. (2011). This is what I'm doing and why: Methodological reflections on a naturalistic think-aloud study of interactive information behaviour. *Information Processing & Management*, *47*(3), 336-348. https://doi.org/10.1016/j.ipm.2010.08.001
Marchionini, G. (2019). Search, sense making and learning: closing gaps. *Information and Learning Sciences*, *120*(1/2), 74-86. https://doi.org/10.1108/ILS-06-2018-0049
Mason, J. (2018). *Qualitative researching* (3rd. ed.). Sage.
McDonald, S., Zhao, T., & Edwards, H. M. (2015). Look Who's Talking: Evaluating the Utility of Interventions During an Interactive Think-Aloud. *Interacting with Computers*, *28*(3), 387-403. https://doi.org/10.1093/iwc/iwv014

Ministry of Education and Research. (2016, 07.06). *Innholdet i ny, femårig mastergrad for grunnskolelærere er fastsatt: Slik blir den nye lærerutdanningen [The content of the new, five-year master's degree for primary school teachers has been determined: This is how the new teacher education program will be]*. Retrieved 27.10.2020 from https://www.regjeringen.no/no/aktuelt/slik-blir-den-nye-larerutdanningen/id2503270/

Mizzaro, S. (1997). Relevance: The whole history. *Journal of the American Society for Information Science*, *48*(9), 810-832. https://doi.org/10.1002/(SICI)1097-4571(199709)48:9<810::AID-ASI6>3.0.CO2-U

Park, T. M. (1993). The Nature of Relevance in Information Retrieval: An Empirical Study. *The Library Quarterly*, *63*(3), 318-351. https://doi.org/10.1086/602592

Pennanen, M., & Vakkari, P. (2003). Students' conceptual structure, search process, and outcome while preparing a research proposal: A longitudinal case study. *Journal of the American Society for Information Science and Technology*, *54*(8), 759-770. https://doi.org/10.1002/asi.10273

Rieh, S. Y., Collins-Thompson, K., Hansen, P., & Lee, H.-J. (2016). Towards searching as a learning process: A review of current perspectives and future directions. *Journal of Information Science*, *42*(1), 19-34. https://doi.org/10.1177/0165551515615841

Robertson, S. (2008). On the history of evaluation in IR. *Journal of Information Science*, *34*(4), 439-456. https://doi.org/10.1177/0165551507086989

Robertson, S. E., & Hancock-Beaulieu, M. M. (1992). On the evaluation of IR systems. *Information Processing & Management*, *28*(4), 457-466. https://doi.org/10.1016/0306-4573(92)90004-J

Sanderson, M. (2010). Test collection based evaluation of information retrieval systems. *Foundations and Trends® in Information Retrieval*, *4*(4), 247-375. https://doi.org/10.1561/1500000009

Sanghee Oh, & Wildemuth, B. M. (2017). Think- aloud Protocols. In B. M. Wildemuth (Ed.), *Applications of Social Research Methods to Questions in Information and Library Science* (2 ed., pp. 198-208). Libraries unlimited. https://ebookcentral-proquest-com.ezproxy.oslomet.no/lib/hioa/reader.action?docID=4743711&ppg=209

Saracevic, T. (1996a). Modeling Interaction in Information Retrieval (IR): A Review and Proposal. *Proceedings of the American Society for Information Science and Technology*, *33*, 3-9. https://www.researchgate.net/publication/239054075_Modeling_interaction_in_information_retrieval_IR_A_review_and_proposal

Saracevic, T. (1996b). Relevance reconsidered. In P. Ingwersen & N. O. Pors (Eds.), *Proceedings of the second conference on conceptions of library and information science (CoLIS 2)* (pp. 201-218). The Royal School of Librarianship. http://tefkos.comminfo.rutgers.edu/CoLIS2_1996.doc

Saracevic, T. (1997). The stratified model of information retrieval interaction: Extension and applications. *Proceedings of the American Society for Information Science and Technology*, *34*, 313-327. http://researchgate.net/profile/Tefko-Saracevic/publication/333293923_The_stratified_model_of_information_retrieval_interaction_Extension_and_applications/links/5ce57ba0458515712ebb7391/The-stratified-model-of-information-retrieval-interaction-Extension-and-applications.pdf

Saracevic, T. (2017). *The notion of relevance in information science: everybody knows what relevance is. But, what is it really?* Morgan & Claypool Publishers.

Savolainen, R. (2018). Pioneering models for information interaction in the context of information seeking and retrieval. *Journal of Documentation*, *74*(5). https://doi.org/10.1108/JD-11-2017-0154

Schamber, L., Eisenberg, M. B., & Nilan, M. S. (1990). A re-examination of relevance: toward a dynamic, situational definition∗. *Information Processing & Management*, *26*(6), 755-776. https://doi.org/10.1016/0306-4573(90)90050-C

Sormunen, E. (2000). *A method for measuring wide range performance of Boolean queries in full-text databases* [Doctoral dissertation, Tampere University]. https://trepo.tuni.fi/bitstream/handle/10024/67002/951-44-4732-8.pdf?sequence=1&isAllowed=y

Soufan, A., Ruthven, I., & Azzopardi, L. (2021). *Untangling the concept of task in information seeking and retrieval* Proceedings of the 2021 ACM SIGIR International Conference on Theory of Information Retrieval, https://pureportal.strath.ac.uk/files/122300382/Soufan_etal_SIGIR_2021_Untangling_the_concept_of_task_in_information_seeking_and_retrieval.pdf

Spink, A., Greisdorf, H., & Bateman, J. (1998). From highly relevant to not relevant: examining different regions of relevance. *Information Processing & Management*, *34*(5), 599-621. https://doi.org/10.1016/S0306-4573(98)00025-9

Strømme, H. (2019). Litteratursøking i kunnskapsbasert praksis og forskning [Literature Searches in Evidence Based Practice and Research]. *Sykepleien*(14). https://doi.org/10.4220/Sykepleienf.2019.61015

Taylor, A. (2012). User relevance criteria choices and the information search process. *Information Processing & Management*, *48*(1), 136-153. https://doi.org/10.1016/j.ipm.2011.04.005

Taylor, A. (2013). Examination of work task and criteria choices for the relevance judgment process. *Journal of Documentation*. https://doi.org/10.1108/JD-12-2011-0054

Taylor, A. R., Cool, C., Belkin, N. J., & Amadio, W. J. (2007). Relationships between categories of relevance criteria and stage in task completion. *Information Processing & Management*, *43*(4), 1071-1084. https://doi.org/10.1016/j.ipm.2006.09.008

Taylor, R. (1968). Question-Negotiation and Information Seeking in Libraries. *College & Research Libraries*, *29*(3), 178-194. https://doi.org/10.5860/crl_29_03_178

Toms, E. G., O'Brien, H. L., Kopak, R., & Freund, L. (2005). Searching for Relevance in the Relevance of Search. In F. Crestani & I. Ruthven (Eds.), *Context: Nature, Impact, and Role. CoLIS 2005. Lecture Notes in Computer Science: 3507* (pp. 59-78). Springer Berlin Heidelberg. https://doi.org/10.1007/11495222_7

Vakkari, P. (2016). Searching as learning: A systematization based on literature. *Journal of Information Science*, *42*(1), 7-18. https://doi.org/10.1177/0165551515615833

Vakkari, P., & Hakala, N. (2000). Changes in relevance criteria and problem stages in task performance. *Journal of Documentation*, *56*(5), 540–562. https://doi.org/10.1108/EUM0000000007127/full/html

Vakkari, P., & Huuskonen, S. (2012). Search effort degrades search output but improves task outcome. *Journal of the American Society for Information Science and Technology*, *63*(4), 657-670. https://doi.org/https://doi.org/10.1002/asi.21683

Vakkari, P., Pennanen, M., & Serola, S. (2003). Changes of search terms and tactics while writing a research proposal: A longitudinal case study. *Information Processing & Management*, *39*(3), 445-463. https://doi.org/10.1016/S0306-4573(02)00031-6

Vakkari, P., Völske, M., Potthast, M., Hagen, M., & Stein, B. (2019). Modeling the usefulness of search results as measured by information use. *Information Processing & Management*, *56*(3), 879-894. https://doi.org/10.1016/j.ipm.2019.02.001

Vanopstal, K., Stichele, R. V., Laureys, G., & Buysschaert, J. (2012). PubMed searches by Dutch-speaking nursing students: The impact of language and system experience. *Journal of the American Society for Information Science and Technology*, *63*(8), 1538-1552. https://doi.org/10.1002/asi.22694

Wacholder, N. (2011). Interactive query formulation. *Annual Review of Information Science and Technology*, *45*(1), 157–196. https://dl.acm.org/doi/abs/10.5555/2766865.2766876

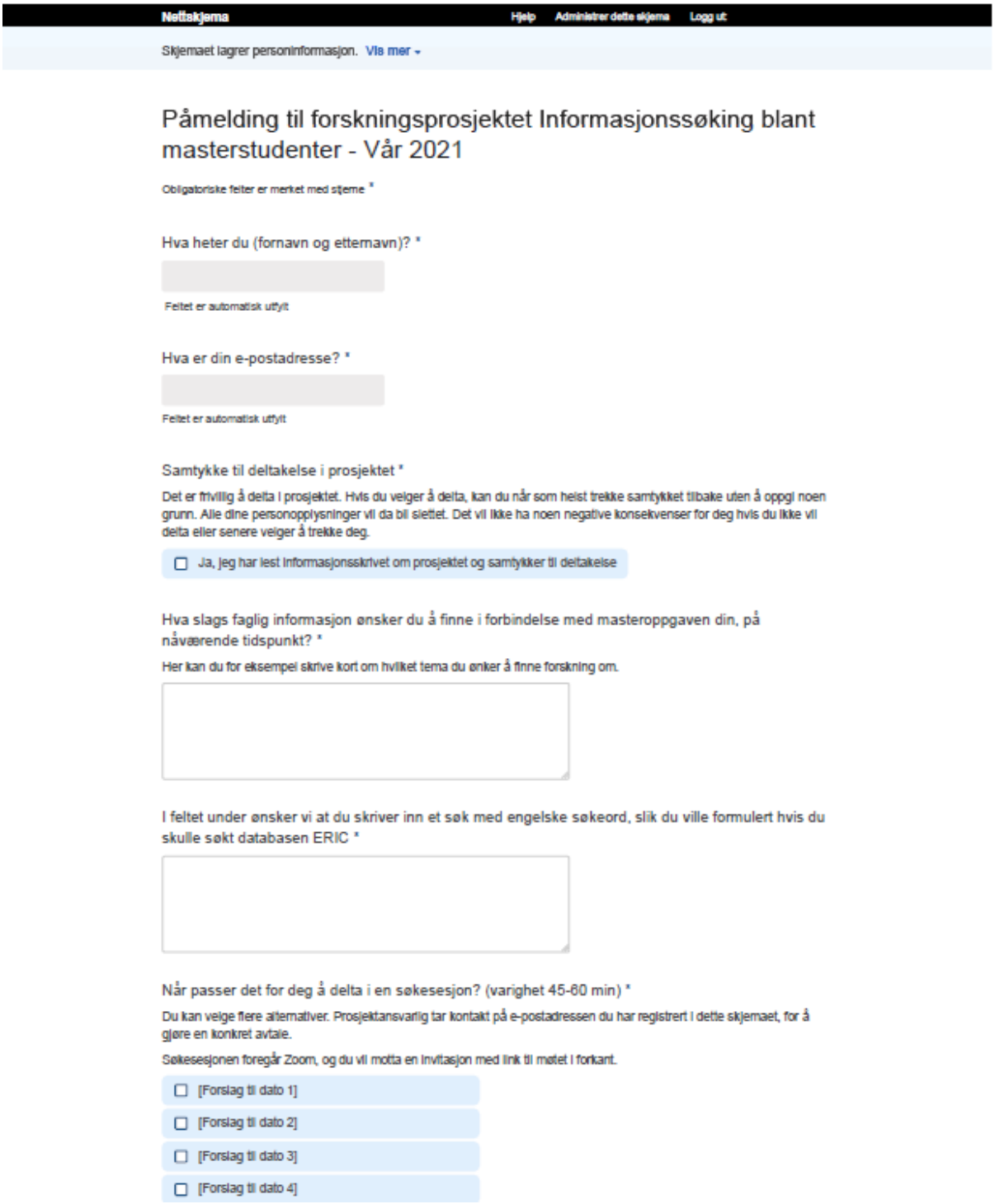
White, R. W., & Roth, R. A. (2009). *Exploratory search : beyond the query-response paradigm*. Morgan & Claypool Publ.

Wilson, T. D. (1999). Models in information behaviour research. *Journal of Documentation*, *55*(3), 249-270. https://doi.org/10.1108/EUM0000000007145

Zhang, Y., & Wildemuth, B. M. (2017). Qualitative Analysis of Content. In B. M. Wildemuth (Ed.), *Applications of social research methods to questions in information and library science* (2 ed., pp. 318-329). Libraries unlimited. https://ebookcentral-proquest-com.ezproxy.oslomet.no/lib/hioa/reader.action?docID=4743711&ppg=329

Østbye, S. E. (2018). *Studenters bruk av Oria. En kvalitativ undersøkelse av informasjonsbehov i møte med et discovery-system* [Master thesis, OsloMet]. https://hdl.handle.net/10642/6723

Appendix A- Nettskjema online recruitment form



Nettskjema Hjelp Administrer dette skjema Logg ut

Skjemaet lagrer personinformasjon. Vis mer

Påmelding til forskningsprosjektet Informasjonssøking blant masterstudenter - Vår 2021

Obligatoriske felter er merket med stjerne *

Hva heter du (fornavn og etternavn)? *

Feltet er automatisk utfylt

Hva er din e-postadresse? *

Feltet er automatisk utfylt

Samtykke til deltakelse i prosjektet *

Det er frivillig å delta i prosjektet. Hvis du velger å delta, kan du når som helst trekke samtykket tilbake uten å oppgi noen grunn. Alle dine personopplysninger vil da bli slettet. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg.

☐ Ja, jeg har lest informasjonsskrivet om prosjektet og samtykker til deltakelse

Hva slags faglig informasjon ønsker du å finne i forbindelse med masteroppgaven din, på nåværende tidspunkt? *

Her kan du for eksempel skrive kort om hvilket tema du ønker å finne forskning om.

I feltet under ønsker vi at du skriver inn et søk med engelske søkeord, slik du ville formulert hvis du skulle søkt databasen ERIC *

Når passer det for deg å delta i en søkesesjon? (varighet 45-60 min) *

Du kan velge flere alternativer. Prosjektansvarlig tar kontakt på e-postadressen du har registrert i dette skjemaet, for å gjøre en konkret avtale.

Søkesesjonen foregår Zoom, og du vil motta en invitasjon med link til møtet i forkant.

☐ [Forslag til dato 1]

☐ [Forslag til dato 2]

☐ [Forslag til dato 3]

☐ [Forslag til dato 4]

Appendix B - Study protocol, pilot tests on Zoom

The following protocol describes a step-by-step account of the procedures in the pilot search sessions performed via Zoom.

Preparation:

1. After receiving participation via Nettskjema and scheduling the appointment, send email to test participant containing:
    a. Confirmation of scheduled date and time
    b. Link to Zoom-meeting

Before the data collection session starts:

1. Set up Screencast-O-Matic for screen recording
2. Open Word document containing participant no. n's two search strings
3. Set up Nettskjema diktafon for recording on phone
4. Check set of data collection papers:
    a. Interview guide
5. Check that pen is available

Data collection:

1. Welcome the test participant to the digital test room
2. Explain purpose of study
3. Explain to the test participant the overall procedure of testing:
    a. Pre search interview
    b. Two different searches. Questions in a think-aloud fashion, regarding relevance for the first 10 results of both searches.
    c. Details on privacy of data collection even when using phone and screen recording (some test participants may ask).
4. Start Nettskjema diktafon for recording
5. Conduct the pre-search interview
6. Start Screencast-O-Matic for screen recording (Nettskjema diktafon will still be recording)

7. Give control to participant via Zoom. Ask test participant to start searching the EBSCO*host* ERIC database via the search box in the web browser.
8. Conduct think-aloud interview for search no. 1
9. When all the 10 first results are assessed, ask if test participant is ready to do search no. 2
10. Using the last version of the search as the starting point, expand it by adding synonyms and additional terms. When participant agrees, copy the search string into the EBSCO*host* ERIC search box in the web browser and run search.
11. Conduct think-aloud interview for search no. 2
12. When all the 10 first results are assessed, ask if test participant has any other comments or questions
13. Thank the test participant for the participation and end Zoom meeting.
14. End Nettskjema diktafon recording and Screencast-O-Matic screen recording

After data collection:

15. Save log data according to test participant no.
16. Collect all paper sheets

Appendix C - Interview guide used in the pilot tests

Introduction, before search no. 1 (Pre-search interview)

What research theme are you planning to investigate in your master thesis project?

At what level would describe your pre-existing knowledge of the research theme?

What kind of information are you looking for at the present moment?

Do you have prior experience with using Boolean operators when searching?

What do you know from before about Boolean operators such as AND and OR?

I will now kindly ask you to perform search no. 1. On the shared screen you can see the EBSCOhost ERIC in the web browser. Control over this screen is given to you via the remote control-function here in Zoom, upon which you can start to search the database. Please do so in your own manner, until you feel that you are satisfied with the results, or that no additional relevant results will be found in a new search.

(Participant runs search no. 1)

How would you judge the relevance of these results? Could you please give a brief assessment of the first ten results? (Or fewer, depending on the no. of actual results)

(During the assessment, the participant will be asked to give statements regarding what relevance criteria are used, and reasons for why result no. X is relevant/partially relevant/not relevant)

(Participant finishes search no. 1)

Thank you. I will now kindly ask you to perform search no. 2. Please find the second search string in the Word document that is open on the computer and paste this into EBSCOhost ERIC in the web browser. You can then run the search.

(Participant runs search no. 2)

How would you judge the relevance of these results? Could you please give a brief assessment of the first ten results? (Or fewer, depending on the no. of actual results)

(During the assessment, the participant will be asked to give statements regarding what relevance criteria are used, and reasons for why result no. X is relevant/partially relevant/not relevant)

(Participant finishes search no. 2)

Thank you. Have you made any considerations during the assessments, concerning differences between the two searches? Did one seem better than the other did, or were they equally useful or equally useless?

Is there anything you would like to add, or do you have any questions?

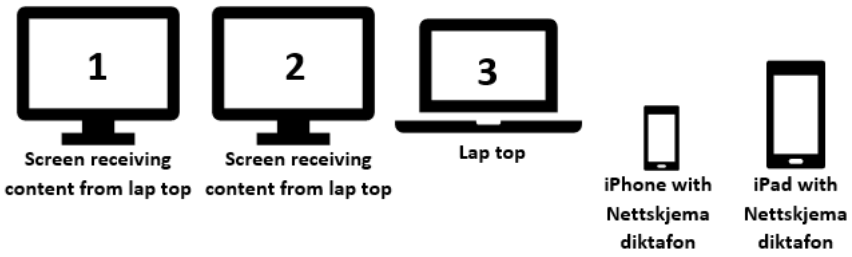
Appendix D - Final study protocol

The following protocol describes a step-by-step account of the procedures of the data collection via Zoom.

Recruitment:

After receiving participation via Nettskjema and scheduling the appointment via email, send a calendar invitation to the test participant, containing a link to the zoom-meeting.

Computer and recorder set up



Before the search session:

1. Start Google Chrome with the EBSCO*host* start page set up for search in ERIC [computer screen 1]
2. Set up Screencast-O-Matic for screen recording of the Google Chrome browser window [computer screen 1]
3. Open Word document containing response from Nettskjema with regards to topic of interest and example of search string, and preconstructed simple and advanced query [computer screen 2]
4. Start Zoom [computer screen 3]
5. Open and prepare Nettskjema diktafon for recording on iPhone and iPad
6. Check set of data collection papers:
    a. Interview guide
7. Check that pen is available

Start search session:

8. Start Zoom-meeting
9. Welcome the test participant to the digital “test room”

10. Explain purpose of study.
11. Explain to the test participant the overall procedure of testing:
    a. Pre-search interview
    b. Two different searches. Questions in a think-aloud fashion, primarily regarding relevance.
    c. Post-search interview
    d. Details on privacy of data collection when using iPhone/iPad and screen recording (some test participants may ask).
12. Start first Nettskjema diktafon recordings on iPhone and iPad
13. Conduct the pre-search interview
14. Finish the pre-search interview and save and send first recording to Nettskjema diktafon server
15. Start second Nettskjema diktafon recording on iPhone and iPad
16. Explain to the participant how the “share screen” and “give control”-functions operate in Zoom.
17. Show and explain the pre-constructed simple Boolean query and give the participant opportunity to comment on the aptness of the query according to information need, and to ask questions. Make changes to the query if needed.
18. Start Screencast-O-Matic screen recording of computer screen 1
19. Give control to participant in Zoom, of Chrome browser window [computer screen 1] and ask him/her to assess the results of the simple Boolean query visible on the screen. Remind participant to think aloud during assessments.
20. Conduct think-aloud interview for search no. 1 – simple query
21. When finished, save, and send second recording to Nettskjema diktafon server
22. Start third Nettskjema diktafon recording on iPhone and iPad
23. Ask if test participant is ready to do talk about the advanced query. Resume control of Chrome browser window [computer screen 1]. Show and explain the pre-constructed advanced Boolean query and give the participant opportunity to comment on the aptness of the query according to information need, and to ask questions. Make changes to the query if needed.

24. Give control to participant in Zoom, of Chrome browser window [computer screen 1] and ask him/her to assess the results of the simple Boolean query visible on the screen. Remind participant to think aloud during assessments.
25. Conduct think-aloud interview for search no. 2 – advanced query
26. When finished, save, and send third recording to Nettskjema diktafon server
27. Stop the Screencast-O-Matic recording and save.
28. Start fourth Nettskjema diktafon recording on iPhone and iPad
29. Conduct post-search interview
30. Finish the post-search interview and save and send fourth recording to Nettskjema diktafon server
31. Thank the participant for his/her time and end the Zoom-meeting.

After testing:

32. Save log data according to test participant number.
33. Collect all paper sheets

Appendix E - Final interview guide

Introduction and pre-search interview

[Participant is welcomed to the digital test room]

Q1 Stated research theme for the master thesis

On what research theme are you planning on writing your master thesis?

Q2 Stated level of familiarity with research theme

At what level would describe your pre-existing knowledge of this research theme?

Q3 Current information need

What kind of information are you looking for at the present moment?

Q4 Intended use of information found in this search

What is the intended use of the information you may find?

Q5 Previous experience searching the ERIC database

Have you used the ERIC database visible on the shared screen, previously?

Q6 Recent experience searching for literature on the stated theme for master the project

Have you searched for information on this research theme recently?

(End of pre-search interview. Next step: An explanation of how the "share screen" and "give control"-functions operate in Zoom is given to the participant. The pre-constructed

simple Boolean query is shown and explained, and the participant is given the opportunity to comment or ask questions. Changes to the query are made if needed. The participant is then given control over the shared screen, and kindly asked to assess the results of the simple Boolean query visible on the screen.)

(Participant starts assessing the results of the simple Boolean query)

Q7 Relevance of results of simple query

How would you judge the relevance of these results? Could you please give an assessment of the results you are seeing on the screen right now?

(During the assessment of results, the participant is asked to give statements regarding what relevance criteria are used, and reasons for the stated relevance.)

(End of search no. 1. Next step: The pre-constructed advanced Boolean query is shown and explained, and the participant is given the opportunity to comment or ask questions. Changes to the query are made if needed. The participant is then given control over the shared screen, and kindly asked to assess the results of the advanced Boolean query visible on the screen.)

(Participant starts assessing the results of the advanced Boolean query)

Q8 Relevance of results of advanced query

How would you judge the relevance of these results? Could you please give an assessment of the results you are seeing on the screen right now?

(During the assessment of results, the participant is asked to give statements regarding what relevance criteria are used, and reasons for the stated relevance.)

(End of search no. 2. Next step: Conduct post-search interview)

Q9 Consideration of both searches

What is your immediate impression of the quality of the two searches? Did one seem better than the other, or did they seem similar?

Q10 Previous experience/knowledge using Boolean operators

How would you describe your previous experiences with using Boolean operators when searching?

Q11 Does the Boolean query seem complex/logical

Do the queries seem complex, in your opinion? Do they seem logical?

Q12 Confidence in using Boolean operators

How would you describe your confidence in using Boolean operators for future searches?

Q13 Likelihood of using Boolean operators

How would you describe the likelihood of you using Boolean operators for future searches?

Q14 Willingness to invest time and effort into learning Boolean

Do you feel it would be worthwhile spending more time learning how to use Boolean operators?

Q15 Further comments

Do you have anything you would like to add, or do you have any questions?

[Participant is given thanks, and the meeting is ended.]

Appendix F - Overview of participants

Participant	Information need and intended use	Simple Boolean query 31 results (peer review filter on)	Advanced Boolean query 52 results (peer review filter on)
P1	*Information need:* Definitions of the concept of «play», and peer reviewed research containing examples of play-based learning in primary education (grades 1 and 2). *Intended use:* To gain overview and knowledge on chosen research theme before defining a research question for the master thesis.	"play based learning" AND "primary education"	("play based learning" or "play-based learning" OR "learn* through play*) AND ("primary education" OR "elementary education" OR "primary school*" OR "elementary school*")
Participant	**Information need and intended use**	**Simple Boolean query** 3 results	**Advanced Boolean query** 29 results
P2	*Information need:* Research that shows how adolescents learn English outside the classroom setting in their spare time, and examples of how this is integrated by teachers in formal educational settings. *Intended use:* To position participant's master thesis in context of earlier research	"language learning" AND "Spare time" AND youth	("Language learning" OR "Second language learning" OR "Foreign language learning" OR EFL OR «english as a foreign language») AND ("Outside the classroom" OR "Self-regulated learning" OR "Autonomous learning" OR "Leisure time" OR "Spare time") AND (adolescen* or youth or teenage*)
Participant	**Information need and intended use**	**Simple Boolean query**	**Advanced Boolean query**

		441 results	561 results
P3	*Information need:* Research or literature on the topic of students' life skills, and on students' perspectives on mental health in school, limited to the senior years of primary education (7th grade in the Norwegian school system) *Intended use:* Gain knowledge on the topic, and utilize in master thesis and a course running parallel to the master thesis course in the fall semester	"life skills" AND "primary education"	("life skills" OR "psychosocial skills" OR «daily living skills») AND ("primary education" OR "elementary education" OR "primary school*" OR "elementary school*")
Participant	**Information need and intended use**	**Simple Boolean query** 14 results	**Advanced Boolean query** 72 results
P4	*Information need:* Research on the topic of minority students and their attitudes towards their own first language. How minority students' first language can be utilized by teachers in the English classroom. *Intended use:* To position participant's master thesis and add precision to research questions and direction, in context of earlier research	"ethnic minority students" AND «first language»	("ethnic minority student*" OR («ethnic background» AND student*) OR "ethnic minority pupil*" OR «ethnic background» AND pupil*)) AND («first language» OR L1 OR «mother tongue» OR «native language»)
Participant	**Information need and intended use**	**Simple Boolean query** 576 results	**Advanced Boolean query** 5711 results
P5	*Information need:* Literature describing or examining the concept of oral communication in textbooks *Intended use:* Inform P5's own methodology and theoretical basis for the master thesis	«oral communication» AND textbooks	("oral communication» OR «Speech skill*" OR "listening skill*") AND (textbook* OR curricul*)
Participant	**Information need and intended use**	**Simple Boolean query**	**Advanced Boolean query**

		18 results	38 results
P6	*Information need:* Any sources on representation of women in literature, including research on women representation in textbooks *Intended use:* To gain oversight and knowledge on chosen research theme	“gender representation” AND textbooks	("gender representation*" OR "wom?n representation*" OR "female representation*") AND (textbook* OR curricul*)
Participant	**Information need and intended use**	**Simple Boolean query** 7 results	**Advanced Boolean query** 230 results
P7	*Information need:* Research on the topic of morning circle groups *Intended use:* Inform theoretical foundation and methodology in the thesis	"classroom community" AND "primary school"	"classroom communit*" AND ("primary education" OR "elementary education" OR "primary school*" OR "elementary school*")

Appendix G – Code book used in NVivo

Codes – inductive analysis

Code	Description
Relevance clues and associated criteria	
Human characteristics	
Affective match - emotional responses to information, fun, frustration, uncertainty	
Curiosity	Participant stating or signifying that he/she became curious as a result of assessment
Empathy	Participant expressing emphatic concern
Enthusiasm	Enthusiastic response during assessment
Fun	Fun as an emotional response to the assessment of an information object
Uncertainty	Uncertainty as an emotional response to the assessment of an information object
Belief match - personal credence given to information, confidence	
Personal credence given to information	Operationalized as it is described
Cognitive match - understanding, novelty, mental effort	
Link to previous knowledge	An information object is linked to the participant's existing knowledge
Novelty	(An information object is assessed as new or original)
Understanding	Assessment of an object triggers understanding
Use or situational match - appropriateness to situation, or tasks, usability, urgency - value in use	
Appropriateness to situation, or task	An assessment of the object as appropriate to the participant's work task
Urgency	An assessment of the object as urgent in terms of usefulness for the participant
Value in use	An assessment of the object as valuable in the context of the participant's work task
Object characteristics	
Content - topic, quality, depth, scope, currency, treatment, clarity	
Certain terms (search via ctrl+f) occuring in full text	Whether or not a certain term is present in the full text document when searched by the participant using ctrl+f
Certain terms (search via ctrl+r) occuring in result list	Whether or not a certain term is present in the result list when searched by the participant using ctrl+f.
Conclusions	The conclusion section of a full text document
Currency (publication date or year)	The publication date or year of a document
Currency not important	The publication date or year stated as not important by a participant
Findings	The findings or results section of an abstract or full text version
Geographic location of study	The geographic location of the study being assessed
Methodology	Type of study or choice of methodology

Scope or depth	The extent of treatment or depth/focus in a study. Can be in terms of age of population in study, a span of years for a set of examined documents in a study or similar
Sources referenced	The sources referenced in the document
Topic	The topic or subject of the document
Treatment	They way in which the topic or subject of a document is being discussed, perspective on topic as expressed in document
Object - characteristics of information objects, e.g., type, organization, representation, format, availability, accessibility, costs	
Abstract	The abstract section of a document or bibliographic reference
Actively not reading abstract	The abstract stated as not important/of little utility by a participant
Recognizes abstract from simple query	
Availability	The possibility of gaining access to a full text version of a document
Full text	The full text version of a document
Full text language	The language in a full text version of a document
Language of bibliographic reference	The language of bibliographic reference
Publication type	Publication type: chapter, article, book etc.
Rank in result list	The rank of a certain result within a result list
Subject terms	The subject terms used to describe a result
Title	The title of a document
Length of title	The number of words in a title
Title recognized from simple query	
Validity - accuracy of information provided, authority, trustworthiness of sources, verifiability	
Author	The author or authors of a document
Journal	The journal of a document
Peer Review	Wheter or not the document is peer reviewed
Publisher	The publisher of a document or journal

Codes – inductive analysis

Code	Description
Challenges with using english terms in search	
Choice of search terms	
Comments on research setting (Zoom, shared screen etc)	
Interaction via shared screen	
Sound quality	
Comments on search habits	
Translating services	
Database functionality or interface in ERIC	
Differences between simple and advanced query	

Familiarity with master thesis topic	
Intended use of information found during search sessions	
Likelihood of using Boolean queries in the future	
Master thesis topic	
Number of retreived results	
Perception of complexity of Boolean queries	
Previous experience using Boolean operators	
Previous experience with ERIC database	
Previous searches on topic by participant	
Relevance assessment	
Not relevant	
Partially relevant	
Relevant	
Stated information need	

Appendix H – NSD Approvement

NSD sin vurdering

Skriv ut

Prosjekttittel
Informasjonssøking blant masterstudenter

Referansenummer
262954

Registrert
20.11.2020 av Didrik Telle-Wernersen - s162447@oslomet.no

Behandlingsansvarlig institusjon
OsloMet – storbyuniversitetet / Fakultet for samfunnsvitenskap / Institutt for arkiv, bibliotek- og informasjonsfag

Prosjektansvarlig (vitenskapelig ansatt/veileder eller stipendiat)
Pia Borlund, piabor@oslomet.no, tlf: 67235250

Type prosjekt
Studentprosjekt, masterstudium

Kontaktinformasjon, student
Didrik Telle-Wernersen, didrik.telle@gmail.com, tlf: 92615045

Prosjektperiode
01.01.2021 - 31.12.2021

Status
26.01.2021 - Vurdert

Vurdering (2)

26.01.2021 - Vurdert
NSD har vurdert endringen registrert 19.01.2021.

Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg den 26.01.2021. Behandlingen kan fortsette.

Zoom er lagt til som databehandler i prosjektet. NSD legger til grunn at behandlingen oppfyller kravene til bruk av databehandler, jf. art 28 og 29.

OPPFØLGING AV PROSJEKTET
NSD vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.

Lykke til med prosjektet!

Kontaktperson hos NSD: Kajsa Amundsen
Tlf. Personverntjenester: 55 58 21 17 (tast 1)

23.11.2020 - Vurdert
Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg den 23.11.2020, samt i meldingsdialogen mellom innmelder og NSD. Behandlingen kan starte.

MELD VESENTLIGE ENDRINGER
Dersom det skjer vesentlige endringer i behandlingen av personopplysninger, kan det være nødvendig å melde dette til NSD ved å oppdatere meldeskjemaet. Før du melder inn en endring, oppfordrer vi deg til å lese om hvilke type endringer det er nødvendig å melde:

https://nsd.no/personvernombud/meld_prosjekt/meld_endringer.html

Du må vente på svar fra NSD før endringen gjennomføres.

TYPE OPPLYSNINGER OG VARIGHET
Prosjektet vil behandle alminnelige kategorier av personopplysninger frem til 31.12.2021.
LOVLIG GRUNNLAG

Prosjektet vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Vår vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 og 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse som kan dokumenteres, og som den registrerte kan trekke tilbake.
Lovlig grunnlag for behandlingen vil dermed være den registrertes samtykke, jf. personvernforordningen art. 6 nr. 1 bokstav a.

PERSONVERNPRINSIPPER
NSD vurderer at den planlagte behandlingen av personopplysninger vil følge prinsippene i personvernforordningen om:

- lovlighet, rettferdighet og åpenhet (art. 5.1 a), ved at de registrerte får tilfredsstillende informasjon om og samtykker til behandlingen
- formålsbegrensning (art. 5.1 b), ved at personopplysninger samles inn for spesifikke, uttrykkelig angitte og berettigede formål, og ikke behandles til nye, uforenlige formål
- dataminimering (art. 5.1 c), ved at det kun behandles opplysninger som er adekvate, relevante og nødvendige for formålet med prosjektet
- lagringsbegrensning (art. 5.1 e), ved at personopplysningene ikke lagres lengre enn nødvendig for å oppfylle formålet

DE REGISTRERTES RETTIGHETER
Så lenge de registrerte kan identifiseres i datamaterialet vil de ha følgende rettigheter: åpenhet (art. 12), informasjon (art. 13), innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18), underretning (art. 19), dataportabilitet (art. 20).

NSD vurderer at informasjonen om behandlingen som de registrerte vil motta oppfyller lovens krav til form og innhold, jf. art. 12.1 og art. 13.

Vi minner om at hvis en registrert tar kontakt om sine rettigheter, har behandlingsansvarlig institusjon plikt til å svare innen en måned.

FØLG DIN INSTITUSJONS RETNINGSLINJER
NSD legger til grunn at behandlingen oppfyller kravene i personvernforordningen om riktighet (art. 5.1 d), integritet og konfidensialitet (art. 5.1. f) og sikkerhet (art. 32).

Nettskjema UiO er databehandler i prosjektet. NSD legger til grunn at behandlingen oppfyller kravene til bruk av databehandler, jf. art 28 og 29.

For å forsikre dere om at kravene oppfylles, må dere følge interne retningslinjer og/eller rådføre dere med behandlingsansvarlig institusjon.

OPPFØLGING AV PROSJEKTET
NSD vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.

Lykke til med prosjektet!

Kontaktperson NSD: Kajsa Amundsen
Tlf. Personverntjenester: 55 58 21 17 (tast 1)