Studying Human Judgments of Relevance: Interactions in Context

Theresa Dirndorfer Anderson

Information and Knowledge Management Program
Faculty of Humanities and Social Sciences
University of Technology, Sydney (UTS)
P.O. Box 123,
Broadway, NSW 2007
Australia
theresa.anderson@uts.edu.au

ABSTRACT

This paper discusses the ways participants in a two-year ethnographic study judged relevance when engaged in searching and research tasks. Two experienced academics have been observed evaluating informative artefacts (documents, citations or other representations) encountered in the course of their own research projects. This study sought to explore the criteria and clues used to make decisions about the relevance of retrievable items. In presenting some of the findings from this longitudinal study, the paper demonstrates the value of this approach for enhancing our understanding of the evolving nature of human relevance judgments. The paper will describe how this interaction involves not only the notion of searcher-system communication, but a range of encounters that inform and influence that particular communication at the search interface. The paper suggests future collaboration between system specialists and human behaviour specialists to further our understanding of the socio-material systems in which people make judgments of relevance.

Categories and Subject Descriptors

H.3.3 Information Search and Retrieval:- search process

General Terms

Measurement, Human Factors

Keywords

human relevance judgments, nature of relevance in contexts, task-based information behaviour

1. INTRODUCTION

Effectively representing information in digital worlds requires us to extend our understanding of the way that searchers in a range of contexts use such representations. Research has demonstrated that, when interacting with an information retrieval system, a searcher's judgements of relevance may not match system determinations of relevance [see for example discussions in: 1, 2, 3]. For a searcher, determining the relevance of the information at hand is fundamental for finding the appropriate information to resolve problems and fill gaps from the overwhelming volume of what is available. A richer understanding of the dynamic human processes associated with assessing the relevance of informative artefacts (documents, citations or other representations) emerges through investigating the way relevance is judged and communicated in authentic, work-based situations. This study sought to explore the interaction between an individual's understanding of their topic (as articulated by them at different points during their research project) and the criteria and clues used to make decisions about the relevance of retrievable items. By investigating the way relevance is experienced by people in the course of their search and research activities, the study sought to provide a detailed account of the evolving and embodied character of human relevance judgments in a manner that could contribute to the development of more interactive, context-sensitive retrieval systems. In presenting some of the richness afforded by this longitudinal process-oriented approach, the paper describes how this interaction involves not only the notion of searcher-system communication, but a range of encounters that inform and influence that particular communication at the search interface.

2. CONCEPTUAL FRAMEWORK

Relevance is a central but much debated concept in Information Science. After more than fifty years of discussion, there is still a lack of consensus about the meaning of the concept. As a result, a number of researchers over the years have observed that, despite its significance, it remains one of the least understood concepts of information retrieval [e.g. 3, 4, 5, 6].

Considerable effort has gone into finding ways to relate searcher perspectives of relevance to those embodied in information systems. From a searcher's perspective, relevance assessment is a process by which she constantly shapes, defines and refines searching. This view of relevance is not contained in a system-

based definition focussing on topic matching between requests and system contents. Over time what has emerged is an acknowledgment that from system and searcher perspectives, relevance is complex, multidimensional and situated, and that the situation of the information user has to be examined in a range of contexts alongside her information need. While it is beyond the scope of this paper to elaborate fully on this extensive body of literature, key outcomes of interest for this study are presented in the next section.

2.1 Overview of Key Assumptions Informing this Study

Process-oriented approaches to studying human judgements of relevance [e.g.:7, 8-11] have explored the varied factors contributing to judgments made by the searchers involved in those studies. A key outcome from such research is the recognition that human judgments of relevance are dynamic and evolving behaviours characterised by the following significant features:

- assessments, even those of subject experts, vary widely and may not be observable in the stated information request [5, 10, 12].
- judgements are related to problem stages, problem solving and work goals [8, 13, 14]; and as a focus forms, clearer criteria for relevance also appear to emerge [15, 16].
- assessment is the result of personal perception and contextual factors, like time, stages of the search process and the context of the information problem [7, 10, 17-19].
- the socio-cultural context of information seeking implicitly establishes criteria for judging the relevance of information encountered during searching and retrieval practices [15, 20].
- while the most likely criteria contributing to relevance judgments are those associated with the information content of documents [14, 21-23], such topically relevant documents are assessed using additional criteria related to the searcher's situation [3, 9, 10, 21].

From this research, we can conclude that perceptions of a search topic and the content of documents change during the course of a research project and therefore investigating a relevance judgment made at any point must be done in context.

Studies like those indicated above demonstrated that human judgments of relevance are based upon criteria extending beyond topic matching between a document and a query. Kekäläinen and Järvelin [24] refer to relevance that is not solely based on topicality as *higher-order relevance* and argue strongly that it plays a clear role in information retrieval interaction, which in turn means *higher-order relevance* must be incorporated into information retrieval system evaluation. Doing so, however, requires a more complete understanding of searcher judgements and the qualities of *higher-order relevance*. Cool [25] and Hert [26] describe these factors and goals associated with retrieval practices as part of the searcher's *situation*. While we have evidence of the multidimensional character of relevance assessment and a range of criteria which may be applied by searchers, we are still unable to ascertain which criteria become most important for users and in which situations this significance might occur. This conclusion suggests an important area for further research, and prompted the study described in this paper.

The *interactive information retrieval* perspective recognises that people use a variety of information sources to learn about a subject or resolve an information problem [13, 18]. As Vakkari [18] observes, viewing information retrieval as interaction between texts and a searcher within a single search session does not allow examination of the whole process of problem solving that generates the search in the first place. This assertion provided another critical thread for the study reported in this paper, which sought to explore relevance assessments as part of the decision-making processes of individuals doing research. Furthermore, it was decided that the searcher must not only be the centre of investigations into relevance, but needs to be allowed to drive the exploration.

The philosophical and methodological position on interaction and context taken in this study draws on the theoretical work of Hert [26], Lave [27] and Barad [28], framing both the system and the searcher positions as complex and dynamic. The significance of these theorists is discussed in fuller detail elsewhere [29]; they are briefly mentioned here to highlight their contribution to both the conceptual and methodological frameworks of this project. Hert [26] argues for a naturalistic, process-oriented investigation to enrich established understandings of information retrieval issues. She depicts searchers working simultaneously in different levels of time-space. Such research demonstrates the complexity and dynamism of the human judgments associated with locating and using information. It also alerts us to the fact that there are no straightforward explanations for the way judgments of relevance are made. Like Hert, Lave's work draws attention to the transformative qualities of human activity and its "embodied, inescapably 'located' nature" [27]. This study draws on Lave's portrayal of both the social and the material structuring of specifically situated activity systems, to frame a searcher's experience of information retrieval systems. Barad [28] takes this idea of socio-material interaction further with her description of *intra-action*.

2.2 Relevance Extending Beyond Traditional Topicality

There is extensive evidence for the dynamic, multidimensional character of relevance [e.g.: 1, 2, 17, 30]. Researchers theorising about relevance have come to appreciate that many of these manifestations of relevance exist side-by-side. The dynamic, multi-dimensional view of relevance described by Saracevic [31], Harter [5] and Greisdorf [4], for instance, draws on the work of Schutz [32] and Sperber & Wilson [33], who use the concept of relevance to explain the complexities of human interactions. For these theorists, relevance is an interacting system of multiple, interdependent relevances. Recognition of such multi-dimensional character of relevance has led to further discussion of the "system of relevances" drawing on Saracevic's earlier references to Schutz and the notion of interdependent relevances [e.g. 4, 20]. Like Saracevic, Mizzaro [6, 17] argues that there are many kinds of relevance. He portrays relevance as having four dimensions: information resources; problem representations; judgment of relevance; and context of the judgment.

While these portrayals used many different terms to describe the elements, it is clear that relevance involves a judgment of some kind. Harter's [5] concept of psychological relevance frames that judgment of relevance as a dynamic mental act. Movement and action are implicit in this notion which, given the situated

character of these human judgments, presses upon us the need to study the human processes used for judging relevance in situations driven by the searcher. In order to do so, relevance has to be examined in the context of everyday practice. Harter [5] and Saracevic [31] also speak of intuitive meanings of relevance in an everyday context, pointing out that we apply it effortlessly when using information, without having to define what relevance is. Schamber [3] goes so far as to suggest that addressing the "everyday, intuitive meaning" of relevance is possible only in a holistic, situational perspective to information behaviour.

Such an everyday sense of relevance is much more complex than topicality as it traditionally defined [5, 34]. Froehlich and Harter argue that relevance judgements made by searchers themselves do not conform to the distinctions made by professional research between different types of relevance. Relevance is the process by which encounters with new information are related to what is already familiar [33]. A phenomena that is relevant to "a matter at hand" – as relevance may be used in everyday language – changes the matter in some way by adding or decreasing information, offering a new perspective, or causing other kinds of cognitive change. In some instances, topicality is not as relevant as other criteria. References on the topic may be less important than references that allow the user to make new intellectual connections. Such a reference is relevant not because it is on the topic, but because it causes a cognitive change [5].

Furthermore, Bean and Green [35] remind us that searchers can use relevance judgments in very different ways: to cast a wide net to trawl through vast amounts of information, but also to narrow, filter and refine. Notions of the concept of relevance can thus appear contradictory or inconsistent to an observer who is not aware of the context in which relevance assessment is made. For this reason, the searcher cannot be considered in isolation from the particular situation in which information is pursued, evaluated and utilised.

Given the diversity of perspectives presented above, any one view of relevance would be incomplete. The research described in the previous section supports the contention that relevance judgements based on topicality (as traditionally defined) alone cannot convey multiple factors underlying a user decision to pursue or use information. Today, building on that earlier work, relevance is recognised as a multilevel phenomenon [24].

From a user's perspective, the judgment of relevance to a subject, or topic, involves more than straightforward matching of terms. Nevertheless, as was noted in the previous section, user-derived criteria show that subject content remains an important factor for relevance – however it is operationalised. If we think of topicality as represented by the search statement (essentially its operational definition), then many user-defined criteria become conditions of selection which do not necessarily identify other criteria, but clarify the user's intended search request. Thus, these other criteria help to express what "on the topic" might mean for the user at a point in time. Furthermore, despite its limits, topicality (relevance to a subject) remains one of the most predominant characterisations of relevance associated with information retrieval. While it is not the only dimension of relevance to consider, topicality is a very practical aspect of relevance because of its operational applicability, observability and measurability [2]. Boyce described it as "operationally necessary" [36].

Perhaps a view of topicality as "subject content" is too narrow to be useful and is not consistent with the intentions of early research. The question statement presented to an information retrieval system contains more than a subject category; it presents a context for those categories. User-centred criteria and qualities related to user selection, satisfaction or evaluation help to clarify the searcher's view of the topic being explored. Hjørland & Christensen [15] support the claim that the expression of a topic is not limited to one representation. The complexity of representations of a topic and searcher expressions of their topic priorities during the course of a search need further exploration.

This growing body of research demonstrates acceptance for the depiction of a dynamic, interacting system of relevances in information retrieval research. While there is still no consensus about the structure of these manifestations, there does appear to be agreement that - ideally, at least - no dimension can be excluded or examined in isolation from the others. Borlund [1] points out that no matter how the classes or types of relevance are explained, judging relevance is based on various criteria. Furthermore, as is discussed in greater detail elsewhere [37], this judgment process has an evolving quality that unfolds through activity – performing a search, making judgements, using representations of information.

3. THE STUDY: BACKGROUND AND METHOD

The study examined relevance assessment within the context of an ethnographic exploration of the research practices of two academics, looking at the ways they experience the concept of "relevance" while using networked information systems.[1] For more than two years, these participants were observed engaged in the organisation, discovery, evaluation and retrieval of information as part of their research practices. Framing the development of the user construct of *topic* in this broader interactional context is consistent with the methodological principles discussed by Denzin [38] and Yin [39].

The ethnographic design of this project followed in the tradition of the user-centred studies of relevance assessment [e.g.: 9, 10, 11] taking their naturalistic, process-oriented principles to the next stage of a micro-level exploration of the dynamism described in earlier research. In particular, the study sought to explore:

- How searchers use informative artefacts (such as documents, citations or other representations) to identify information that is *relevant* to them; and
- How the meaning of relevance in relation to the *topic* prompting search or research activity is communicated at various stages of a project.

The starting point was searchers using networked information systems, but the searchers' individual research interests, reactions and responses to information they encountered during the course of engagement with them drove the inquiry. These participants were observed searching and evaluating both networked and print

[1] In this study networked information systems were defined as all information systems accessible to participants from their desktop computers. This included all databases available via their university library as well as web-based resources such as websites, digital libraries & internet-based databases.

information resources and formats of citations, abstracts and texts. They were also observed preparing documents as part of their research work. The fieldwork thus used process-oriented methods of discovery allowing participants to shape the inquiry and for information seeking to be observed in context – that is, as part of their own ongoing research projects.

Both participants ("Catherine" and "John"[2]) were experienced users of networked information systems who were also experienced academics at, or near, the beginning of research projects that involved the use of networked information systems. Fieldwork involved engaging as a "participant-observer" with them, observing and discussing their discovery, selection, evaluation and use of information. To more fully understand how each judged relevance during their search practices, it also involved examining expressions of their topic and the processes by which they made sense of what they found. A multi-layered narrative was created by weaving together different ethnographic stories - *impressionist tales* [40] along with anecdotes and vignettes [41, 42] - with passages from field notes, email correspondence, video and audio records and other documents associated with the *story* of the two participants. An extract from John's story is provided in the appendix (*Triggers Working Together*).

Inductive analysis was carried out on the fieldwork to gain a rich understanding of the way these researchers worked with informative artefacts to evaluate the relevance of information they encountered throughout the two-year study. Ethnographic storytelling [43, 44] and narrative inquiry [45] served as tools for both the analysis and presentation of this process. In other words, the production of narratives about Catherine's and John's experiences became an integral part of understanding and representing what was observed. At the same time, the analysis involved a distillation of the texts and stories produced in the process so that, like the layers of an onion peeled away, core observations could be selected for presentation. The overall focus of the stories was a description of the process by which Catherine and John discovered meaning in relation to their search and research goals. The analytical themes used to select elements from field texts can be summarised as:

- Expressions of a "topic" within the context of an utterance, statement or observed action;
- Communication of the expression to mechanical (e.g.: database, search engine) and human "retrieval systems" (e.g.: research assistants);
- Evaluation of "relevant" responses resulting from interaction with those human and mechanical systems.

The layered transcriptions created for each recorded encounter with the participants are at the heart of these narrative forms. Texts of the audio and video recordings of search sessions, relevance evaluations and discussions were created through repeated listening and watching of each recording. These texts combine words and actions observed on tape with field notes and analysis prompted by hearing comments and watching actions in the context of the recorded event. In this way, the writing of the research narrative became a powerful device for understanding relevance interactions.

4. FINDINGS: TRIGGERS, CLUES AND CRITERIA FOR JUDGING RELEVANCE

The field research contributed to an extended analysis of the evolving character of participants' judgements of relevance and of the topics they were exploring in their own research projects. One critical component of the study reported here involved examining how Catherine and John used informative artefacts (e.g.: database citations, abstracts, document records.) to determine the relevance of the information represented by such artefacts.

4.1 Triggers Signalling Action

Catherine and John looked for clues to help them determine content, applying their knowledge of authors, journals and genres to judgements about documents (or representations of documents) they were examining. In addition, they have been observed looking for and making use of "trigger words" in citations, documents or referential material associated with an item. John used this term throughout his involvement in the study to describe words or phrases that prompted him to act in relation to the document or citation at hand. His use of the term indicated that he was making a distinction between triggers and clues. This finding prompted a deeper examination of the material collected in the field and showed that Catherine described similar behaviour, although she did not use the term "trigger words." More than simply providing a clue about the content, a trigger encouraged action.

Triggers could lead to the selection but also the de-selection of information. Decisions to select or reject articles were based on a combination of factors that worked together in varied ways. These criteria were not mechanistic devices applied to their choices. Sometimes they were able to make a quick decision one way or the other, but other citations took careful reading and reflection before a decision was made. Specifically, analysis of the participants' evaluation processes builds on earlier relevance criteria research and suggests:

- **Authorship** triggered relevance judgements by providing clues about the content of a document or representation.
- **Titles** triggered relevance judgments by providing a sense of an article's likely content, helped with selection decisions.
- **Particular genres, or information types,** triggered relevance judgments at the boundaries of a search.
- **Trigger words** aided the selection process, helping to manage the exploration of unfamiliar literature.

Triggers worked together to prompt very personal reactions. Discovering familiar authors shaped the participants' perceptions

[2] The participants (referred to here by their pseudonyms: Catherine and John), were both senior academics at the author's university. They were invited to participate in this project on the recommendation of the university's liaison librarians, who were asked to help identify potential research participants from within the UTS academic community. The aim was to identify academics who were i) in disciplines other than library & information science; ii) in the early stages or about to embark upon a research project that would require them to explore the networked information resources available via the university's online network and iii) were not novice online searchers. Catherine and John were the first recommendations made, and both readily agreed to participate.

of texts or their representations. Spotting favourite writers, significant personalities in the field, familiar names in the author field or in a reference list provided some sense of what the item under review could be "about." Catherine's familiarity with the work of a noted researcher in one of her areas of interest, for instance, enabled her to frame the contents of an article by an unfamiliar author, when a quick scan of that article's research findings and references confirmed that the work of the known personality had provided the framework for that unknown author's own research.

Spotting familiar authors also helped Catherine and John to work out how an item might relate to their own research or search goals. Articles mentioning a familiar or significant author made it hard for John to dismiss them, even if he had reservations about the way the content might relate to his needs. In the first moments of reviewing a series of abstracts printed out as a result of database searching a month earlier, for example, John very quickly pointed out that the author of the first abstract is someone whose work he values highly:

> *I don't know whether that's specifically to do with social evolution. It's just that I'm familiar with and respect her work and so on those grounds alone I'd like to read what's in there ….*

Works by and about writers he respects also acted as triggers for John because such works tend to be "talking about all the right things."

Similarly, Catherine kept an eye out for favourite authors whose work she respected and for writers she recognised as significant personalities in her field. Authorship had an observable impact on her reaction to items and any further action she chose to take with a document. Without even reading an article, Catherine's knowledge about authors and their position in her field often led to certain expectations about content. Her perception of "key writers" allowed her to build a picture of the type of material she expects to see in their work. In such instances, authorship is a critical selection criterion even if she is not intending to follow their research approach. Even without abstracts, Catherine often knew enough about particular citations to select them. Furthermore, when encountering a jointly authored text, recognising one author as being a particular favourite helped her decide if it could be useful.

The name of the journal, like the authorship trigger discussed earlier, helped both participants to situate the content of a document or document representation in relation to their own work and research needs. Mention of a journal like *World Futures*, for instance, signalled John that an item could be useful for his research because it was a journal he respected. Similarly, the title of the article itself facilitated this process. For example, John considered the article "What two legs can learn from four legs" important because the title

> *… [suggests] this link between humans and other mammals… there might be something there.*

However, there were also occasions when John decided **not** to select the citation – even though he did not fully understand the content – because his "sense" about triggers spotted in a text or citation was enough to confirm his decision **not** to select an item. When reviewing the citation "Reality, artifice, and the politics of evolution: Watts and Carlyle in the Earnest Age," John read the title quietly aloud before saying:

> *I haven't a clue what that's about…Victorian Poetry [the journal title] – I'm fairly suspicious that's going to be marginal and I'm prepared to take a risk on that [feeling].*

The journal containing the article, combined with a gut feeling about its marginal connection to his topic, offered the clues he needed to exclude this citation.

Another aspect emerging from the analysis of these experiences was that journals, particularly those titles familiar to a searcher, could serve as valuable tools for managing the scale of a project. John, in particular, showed a preference for book reviews over specific articles or even the original book in instances where he was trying to manage his collection and find a "way in" to a topic. He demonstrated a tendency to select journal articles to focus on specific aspects of his core research objectives. Books and book reviews, on the other hand, were often selected to address peripheral issues. Using book reviews instead of detailed journal articles appeared to be a pragmatic strategy to avoid "taking things too far." Often John was not necessarily interested in getting the books themselves, but relied upon the book reviews to provide different perspectives of the book content and identify valuable parallels and analogies for his work.

Finally, certain words or word combinations acted as "triggers" – helping to flag the potential significance of the texts or representations under review. Trigger words can come from anywhere, as Catherine found out in one series of searches. Attending a conference the day before one of her search sessions sparked the realisation that the terminology she had been trying to use to access information about her topic had changed recently. This shift suggested an alternative to the terms she had used in her earlier searches. The search terms she had entered in observed searches just a few weeks earlier were no longer in common usage. This discovery helped explain why she had not been able to find the information she expected in that first search session. More importantly, it showed her how important it was to have the "right word" when preparing her database searches. Reflective of the iterative character of information seeking and the relevance judgments taking place along the way, the generation of these new search terms emerged through Catherine's engagement with information retrieved through her database searches as well as through personal encounters at the conference. Such practice offers a glimpse of the inter-relation between judgments about the appropriateness of a search statement and the search results themselves.

Speaking with Martin (John's research assistant) about the process he and John used to identify "relevant" items showed that he was also relying on certain words in his evaluations:

> *You very quickly become used to seeing the key words that are popping up all the time.*

He then rattled off a range of terms that jumped out from the texts in his hand at that moment that were similar to the triggers prompting John to placing certain articles in his "highly relevant" collection of articles.

4.2 Clues and Interactions Shape Understanding

The clues and triggers discussed in the previous section were not applied in isolation; instead it was the interplay of criteria associated with the title, the author and personal knowledge of authors (through familiarity with their work and/or personal contact) that provided important clues to the content. Interaction with colleagues and with written texts (in a range of forms) played a pivotal role in the ways John and Catherine formulated their topics. These "texts" include databases, people, and written representations ranging from online citations through to the "full texts" of retrieved material and papers prepared by colleagues, contacts and themselves. Experiences with these other texts help them to frame their own work, articulate their own topics and respond to the content of other texts they wished to evaluate. Even when presented with the full citation of a document, the title of each item appeared to play an important role in "setting the scene" for both of these researchers. When trying to articulate what they were looking for in the document or representation under review, they described connections to other material they had read, people they had encountered, and experiences they recalled.

The extract from John's story (see the appendix) shows how drawing a link between words in the title and *social evolution* as a database keyword signalled John what citations could be interesting for his research. Making such connections creates trigger words providing potential pathways into the literature a few weeks later.

In particular, analysis of both participants' observed experiences (like the one described in the appendix) reveals the following four main features of this interaction.

4.2.1 Finding Frameworks for Research and Search

In conversations with John throughout his involvement in this study he described how email communication and personal meetings with international collaborators was shaping the categories created, the themes to be covered, and the extent of that coverage in his research. The background and particular interests of Martin (John's research assistant) also appeared to influence the evolution of John's research topic. In turn, this inter-relationship impacted upon Martin's own as well as John's interpretations of the selection criteria John had established for the literature under review. Similarly, devising course outlines for the forthcoming semester provided a framework for the research Catherine began to work on at that time. A chance encounter with an informal 8-page document also helped her establish clearer limits for her research, making it easier for her to articulate what she wanted to cover in her research. In this frame of mind, she revisited some of the earlier material she had collected for her research and located a document that provided the framework enabling her to move even further forward and utilise her other material in what she felt was an efficient way.

4.2.2 Awareness of a Relationship to Other Literature Encountered

Interaction with other texts, both those they were creating and those of other writers, affected the participants' relevance judgements and views of their topic. Such encounters provided links that made some citations valuable enough to pursue, whilst signalling that others were inappropriate. Catherine is observed drawing on her existing knowledge (of the author, professional ties and the suggested theme) when evaluating items. Similarly, whilst observing John's evaluations of abstracts in the early months of the study, he identified many links to the themes raised in a particular text that he had been using to prepare a research proposal at that time, prompting him to widen the scope of the literature he wanted to collect. Furthermore, the timing of these encounters is critical. Writing that research proposal and simultaneously reviewing retrieved abstracts helped John spot links between these search results and key features of his project as he was beginning to articulate them. In this way, specific terms emerged as very powerful triggers during those abstract reviews and subsequent evaluation sessions by both John and his assistant.

4.2.3 Triggers Work Together

John categorised his intended use for the articles he was reviewing. During the evaluations depicted in the extract from John's story (see the appendix), for instance, those he selected were relevant in a specific way and for a specific purpose, based on what he knew about the theme presented in the particular abstract he was reviewing at the time. This "knowing" was a product of his interaction with: (i) his established understandings of the topics recognised in this text representation; (ii) the literature he had read before encountering the representation under review (in this instance: discussions of key drivers of evolution); and (iii) the earlier works in this *particular database* collection exploring one or more of these drivers. His reactions to "triggers" help him decide when the risk is worth taking. His understanding of the terminology and perceptions of the type of article that would be contained in the journals helps him in his judgements. Similarly, Catherine's judgements illustrate how the title, the author and her personal knowledge (of authors and their professional affiliations) combine to provide important clues regarding the significance of the content.

4.2.4 Clues about Database Contents Create Expectations

John's preliminary judgements about the "potentially relevant" documents available within each of the databases he searched were influenced by his pre-existing knowledge of the content of and experience with each particular database. This familiarity appeared to influence his selection decisions. Similarly, Catherine's explorations were often guided by what she knew – or had been told by liaison librarians – about the database she was using.

5. DISCUSSION: LEARNING BY APPLYING RELEVANCE CRITERIA

The descriptions of judgments of relevance in this paper portray searchers who are learning about the topics that interest them through the process of engaging with informative artefacts of various kinds. They have been observed looking for clues within such representations of information that will help with the selection as well as the elimination of items under review. The evolving character of these judgments is most effectively observed using a longitudinal process-oriented approach to information retrieval interaction. As has been illustrated in the depictions of searcher experiences in this paper, informative artefacts provide prompts for further action. Often, the item itself is not relevant enough to be selected for further use, but nevertheless can play a critical role in the overall process of a

person's meaning making. Furthermore, this ethnographic approach also demonstrates how human interactions impact judgments of relevance in ways that are often undetectable in studies of searcher-system interactions.

While the behaviours described in the previous section resemble the tactics, stratagems and strategies reported by Bates, for example [46], these findings contribute further understanding about how embedded such practices are in the broad context of the participants' research interests. These findings extend earlier research on the use of topicality as a criterion and the richness of topical relevance [35]. Discussions with Catherine and John and observations of their evaluations of the topic presented in citations or full articles were associated with criteria like author, personal experience, currency, and uniqueness. Interactions with colleagues (through personal contact or encounters at conferences) or written texts (either their own or those prepared by others) produced some of the triggers that helped Catherine and John decipher the content of material they examined in a manner most appropriate to their personal situation. While the study did not seek to identify specific criteria used during evaluations, Catherine and John seemed to refer to non-topic elements of an item as part of its content.

Observations undertaken at various stages of their respective research projects support and extend earlier findings by Bateman [23], who observed a strong preference by non-expert users for making ultimate relevance judgements after reading the information pursued in a search situation. There were many occasions when information considered highly relevant by Catherine and John was identified without looking at the full text of the item. However, there were also instances when such value was confirmed only after reading the text or at the very least skimming its content. This suggests that, in spite of discipline expertise or familiarity of authors and journals, there are still situations when a document representation is inadequate for judgements of relevance.

In summary, the dynamic nature of the interplay between criteria is illustrated throughout Catherine and John's stories. This study also demonstrates the significant impact that context has on this decision-making process. In this way, the portrayal of interaction in this study reflects Barad's [28] notion of intra-action, to emphasise the emergent quality of working with a topic and understandings about what constitutes a relevant piece of information in the course of dealing with information systems (human as well as mechanical), people and informative artefacts. Given the insights afforded by detailed exploration of real-life contexts, the present study can contribute to future design of context-sensitive information systems by providing a rich depiction of the evolving and embedded qualities of human relevance assessment processes.

6. LOOKING TO THE FUTURE

During a discussion with John, he described his approach to his research and how he saw his searching at that point fitting in with his work:

> *"…So it's like kind of building a building from a number of different starting points and I just hope they all integrate."*

John's statement is reflective of the naturalistic, non-sequential character of the research processes observed in this study. It is an equally applicable description of the naturalistic approach taken to study the information practices reported in this paper. This fine-grained approach to examining the nature of human judgments of relevance over time and in the context of a searcher's broader context (the participants' own research projects) contributes to a more complete understanding of the evolving nature of these searcher judgments. By showing how this dynamism is embodied in context, the rich description and interpretation afforded by this study contribute to our growing understanding of higher-order relevance. However, the story can be made even richer by relating these human experiences more substantially to the material structure of the information systems and system interpretations of the human activity described in this study. Thus, this paper is an invitation for further collaboration between system specialists and human behaviour specialists, for example to create richer scenarios of task-based contexts or strategic helps for searchers. In this way we can enhance our understanding of the socio-material systems in which people make judgments about the relevance of information.

7. References

1. Borlund, P., *The concept of relevance in IR.* Journal of the American Society for Information Science and Technology, 2003. 54(10): p. 913-925.
2. Schamber, L., M.B. Eisenberg, and M.S. Nilan, *A re-examination of relevance: Toward a dynamic, situational definition.* Information Processing and Management, 1990. 26(6): p. 755-776.
3. Schamber, L., *Relevance and Information Behavior.* Annual Review of Information Science and Technology, 1994. 29: p. 3-48.
4. Greisdorf, H., *Relevance: An interdisciplinary and information science perspective.* Informing Science, 2000. 3(2): p. 67-71.
5. Harter, S.P., *Psychological relevance and information science.* Journal of the American Society for Information Science, 1992. 43(9): p. 602-615.
6. Mizzaro, S., *Relevance: The Whole History.* Journal of the American Society for Information Science, 1997. 48(9): p. 810-832.
7. Tang, R. and P. Solomon, *Toward an understanding of the dynamics of relevance judgement: an analysis of one person's search behaviour.* Information Processing and Management, 1998. 34(2/3): p. 237-256.
8. Wang, P. and D. Soergel, *A cognitive model of document use during a research project. Study 1. Document selection.* Journal of the American Society for Information Science, 1998. 49(2): p. 115-133.
9. Barry, C.L., *User-defined relevance criteria: an exploratory study.* Journal of the American Society for Information Science, 1994. 45(3): p. 149-159.
10. Park, T.K., *The nature of relevance in information retrieval: An empirical study,* in *School of Library and Information Science*. 1992, Indiana University (UMI No. 9231644): Bloomington, IN. p. 209.
11. Schamber, L., *Users' criteria for evaluation in multimedia information seeking and use situations,* in *Graduate School*. 1991, Syracuse University (UMI No. 9214390): Syracuse, NY. p. 327.
12. Swanson, D.R., *Subjective versus objective relevance in bibliographic retrieval systems.* Library Quarterly, 1986. 56(4): p. 389-398.

13. Byström, K. and K. Järvelin, *Task complexity affects information seeking and use.* Information Processing & Management, 1995. 31(2): p. 191-213.
14. Vakkari, P. and N. Hakala, *Changes in relevance criteria and problem stages in task performance.* Journal of Documentation, 2000. 56(5): p. 540-562.
15. Hjørland, B. and F.S. Christensen, *Work tasks and socio-cognitive relevance: a specific example. Brief Communication.* Journal of the American Society for Information Science, 2002. 53(11): p. 960-965.
16. Vakkari, P., *Task-based information seeking*, in *Annual Review of Information Science and Technology*, B. Cronin, Editor. 2003, Information Today: Medford, NJ. p. 413-464.
17. Mizzaro, S., *How many relevances in information retrieval?* Interacting with Computers, 1998. 10(3): p. 303-320.
18. Vakkari, P., *Task complexity, problem structure and information actions: Integrating studies on information seeking and retrieval.* Information Processing & Management, 1999. 35(6): p. 819-837.
19. Spink, A., H. Greisdorf, and J. Bateman, *From highly relevant to not relevant: examining different regions of relevance.* Information Processing and Management, 1998. 34(5): p. 599-621.
20. Cosijn, E. and P. Ingwersen, *Dimensions of relevance.* Information Processing & Management, 2000. 36: p. 533-550.
21. Barry, C.L. and L. Schamber, *Users' criteria for relevance evaluation: a cross-situational comparison.* Information Processing and Management, 1998. 34(2-3): p. 219-236.
22. Maglaughlin, K.L. and D.H. Sonnenwald, *User perspectives on relevance criteria: A comparison among relevant, partially relevant, and not-relevant judgments.* Journal of the American Society for Information Science and Technology, 2002. 53(5): p. 327-342.
23. Bateman, J., *Changes in relevance criteria: a longitudinal study*, in *Proceedings of the 61st American Society for Information Science Annual Meeting*. 1998. p. 23-32.
24. Kekäläinen, J. and K. Järvelin, *Evaluating information retrieval systems under the challenges of interaction and multidimensional dynamic relevance*, in *CoLIS4 : proceedings of the Fourth International Conference on Conceptions of Library and Information Science, Seattle, WA, USA, July 21-25, 2002*, H. Bruce, et al., Editors. 2002, Libraries Unlimited: Greenwood Village, CO. p. 253-270.
25. Cool, C., *The nature of situation assessment in new information retrieval environments*, in *ASIS '97: Digital Collections: Implications for users, funders, developers and maintainers. Proceedings of the 60th ASIS Annual Meeting Washington, D.C. November 1-6, 1997*, C. Schwartz and M. Rorvig, Editors. 1997, Information Today for ASIS: Medford, NJ. p. 135-146.
26. Hert, C.A., *Understanding information retrieval interactions: theoretical and practical implications*. Contemporary Studies in Information Management, Policy, and Services, ed. P. Hernon. 1997, Greenwich, CN: Ablex.
27. Lave, J., *Cognition in Practice: Mind, mathematics and culture in everyday life*. 1988, Cambridge, UK: Cambridge University Press.
28. Barad, K., *Getting real: technoscientific practices and the materialization of reality.* Differences: A journal of feminist cultural studies, 1998. 10(2): p. 87 (4).
29. Anderson, T.D., *Understandings of relevance and topic as they evolve in the scholarly research process*, in *Faculty of Humanities and Social Sciences*. 2003, University of Technology, Sydney.
30. Saracevic, T., *Interactive models of information retrieval (IR): A review and proposal.* Proceedings of the 59th Annual Meeting of the American Society for Information Science, 1996. 33: p. 3-9.
31. Saracevic, T., *Relevance reconsidered '96*, in *Proceedings of CoLIS 2, second international conference on conceptions of library and information science: Integration in perspective*, P. Ingwersen and N.O. Pors, Editors. 1996, Royal School of Librarianship: Copenhagen. p. 201-218.
32. Schutz, A., *Reflections on the problem of relevance. Edited, annotated, and with an introd. by Richard M. Zaner*. 1970, New Haven, CT & London: Yale University Press. xxiv, 186 p. 21 cm.
33. Sperber, D. and D. Wilson, *Relevance: Communication and cognition*. 1986, Oxford: Basil Blackwell.
34. Froehlich, T.J., *Relevance reconsidered - towards an agenda for the 21st Century: introduction to special topic issue on relevance research.* Journal of the American Society for Information Science, 1994. 45(3): p. 124-133.
35. Bean, C.A. and R. Green, *Relevance relationships*, in *Relationships in the organization of knowledge*, C.A. Bean and R. Green, Editors. 2001, Kluwer Academic: Dordrecht, The Netherlands. p. 115-132.
36. Boyce, B., *Beyond topicality. A two-stage view of relevance and the retrieval process.* Information Processing and Management, 1982. 18: p. 105-189.
37. Anderson, T.D., *Relevance as process: judgements in the context of scholarly research.* Information Research, 2005. 10(2): p. paper 226.
38. Denzin, N.K., *Interpretive interactionism*. Applied Social Research Methods Series, Volume 16. 1989, Newbury Park, CA: Sage.
39. Yin, R.K., *Case study research: design and methods*. 1994, Thousand Oaks, CA: Sage.
40. Van Maanen, J., *Tales of the field: on writing ethnography*. 1988, Chicago: University of Chicago.
41. Ely, M., et al., *Doing qualitative research: circles within circles*. 1991, London: Falmer Press.
42. Ely, M., et al., *On writing qualitative research: living by words*. 1997, London: Falmer Press. 411.
43. Bochner, A.P., *Narrative's virtues.* Qualitative Inquiry, 2001. 7(2): p. 131-157.
44. Richardson, L., *Writing: a method of inquiry*, in *Handbook of qualitative research*, N.K. Denzin and Y.S. Lincoln, Editors. 2000, Sage: Thousand Oaks. p. 923-948.
45. Clandinin, D.J. and F.M. Connelly, *Narrative inquiry: experience and story in qualitative research*. 2000, San Francisco, CA: Jossey-Bass.
46. Bates, M.J., *Where should the person stop and the information search interface start?* Information Processing & Management, 1990. 26(5): p. 575-591.

Appendix: Extract from John's Story: *Triggers working together*

I am watching John as he reviews the abstracts of citations he selected two weeks earlier. He questions the focus of *social evolution* taken in one abstract, uncertain if the term is being used in a

"…focused or considered way. It looks like it could be more general on paper. But at least it's worth having a look at."

He discusses past experiences that are shaping his judgement here. The terms *social evolution* and *evolution in society* are sometimes used in superficial ways, while

"…other times it's a very meaningful and considered statement."

He illustrates this point by talking about his recent experience with a book titled *Sensitive Chaos* whose very title enticed him to skim through its pages. It turns out that the book's focus is interesting in its own right, but not in terms of John's study. It proved to be a disappointment because:

"it didn't have any real link in with social evolution in the way I'm interested."

With this recent experience still fresh in his mind, abstracts like the one he just read can pose a problem. Sometimes it seems as if every abstract reveals yet another relationship or association within the domain of *social evolution* that John's project will ultimately have to address. John refers to them as "nodes" - term he often uses to describe all the threads of his topic that seem to be emerging during this search process. In the abstract for a book about Habermas and the politics of discourse, for example, he sees

"…the link between social evolution and morality which brings us back to the ethical literature."

More than that, this work marks the formation of yet another area that John's project will ultimately have to address.

"… All it is doing is identifying a body of knowledge that I need to be familiar with and then strategically I have to work out how I can do that with the least investment of time and energy."

This begs me to ask him how he foresees doing that at the moment. He laughs before adding:

"Well, clearly the first thing to do is develop a familiarity with the relevant literature and see whether in that literature there seem to be key works that will get me sooner than later a familiarity with that literature sufficient for my wider purpose. So, there's a kind of strategic approach here."

As he makes this observation, he points to the citation indicators at the top of the reference that he has just noticed and which identify this item as a book abstract, not an article as he originally thought. It is hardly surprising that he should be confused by the abstract. John tends to use this particular database to locate serials literature. While he has encountered many book reviews and review articles today, this is the first abstract for a book he has recently read. He is also trying to move quickly through his review, because he has to leave soon to get on with class preparations. It is easy to miss the "Book abstract" identifier at the top of this citation. Now that he has, it seems easier to work out his next step:

"…so it might be that by getting that book I could very quickly get a road map for the big area and the critical texts -- which I need to do."

With that he moves on to review the next citation abstract. A brief glance at an item tells him this could be a "fairly specific article" that could still be very useful because:

"It's got all the triggers we've discussed already – or several of them."

However, the next abstract doesn't contain the right triggers even though a month ago, when he initially selected this citation for further evaluation, he thought it looked promising. His voice trails off, as he points to terms in the abstract. He pauses for a moment as he ponders the potential of this text.

The abstracts John has been selecting for further review have been relevant for a range of reasons. He talks about the risk he takes excluding or ignoring information that could be valuable for his research. An item does not have to be selected to be useful. Sometimes the records he is reviewing merely have to help him develop a better understanding of the concept of *social evolution*. Once they have served that purpose, he no longer needs to pursue the article itself. His reactions to "triggers" help him decide when the risk is worth taking.