

Injecting Realism into Simulated Work Tasks: A Case Study of the Book Domain

Pia Borlund

Oslo Metropolitan University, Norway, piabor@oslomet.no

Toine Bogers

Aalborg University, Denmark, toine@hum.aau.dk

ABSTRACT

Interactive IR studies are commonly conducted by assigning participants *simulated work tasks situations* (SWTS), predetermined textual descriptions of an information need and its context to motivate participants to perform realistic searches of the IR system in question. Despite their popularity, research has shown that formulating realistic and engaging SWTSs remains challenging. In this paper, we propose using the analyses of large samples of complex real-world search requests—with the specific case of book search—to inject realism into SWTSs by taking different relevance aspects and their relative frequencies into account when formulating them. This should make it easier for researchers to design realistic SWTSs in the future.

KEYWORDS

Simulated work tasks, interactive IR, complex search, book search

INTRODUCTION

In interactive information retrieval (IIR) studies, participants are commonly assigned predetermined search tasks to initiate their search behavior. These search tasks function as a test instrument enabling realistic search interaction and assessment of retrieved information. As a result, it is necessary for these tasks to be designed in such a manner that they fit the study participants, so that they can adopt these search tasks as if they were their own personal information needs. This requires search tasks to be designed with great care to ensure their soundness and validity. Anybody who has conducted an IIR study using assigned search tasks (e.g., in the form of *simulated work task situations* (Borlund, 2000; 2003)) will recognize that this is straightforward in theory, yet difficult in practice. Research shows that the majority of IIR studies fail to sufficiently tailor the simulated work task situations (SWTS), despite being rich in contextual descriptions, hence also disqualifying the study results (Borlund, 2016).

We believe that the design of SWTSs could benefit from a better understanding of real-world search tasks and information needs, and there is rich history of studying this. Recent work by Koolen et al. (2015) and Bogers et al. (2018) has taken a different approach to studying real-world needs by analyzing the complex requests of users in book and movie discovery. Their focus was on the relevance aspects expressed by users in their complex requests, i.e., what makes a book or movie relevant to a person's request? The objective of the paper is to help IIR researchers design better and more realistic search tasks for IIR evaluation studies—more specifically book search tasks—by using the relevance aspects uncovered by Koolen et al. (2015) to serve as a source of inspiration for the design of SWTSs. In the next section, we revisit the general evaluation framework of SWTSs, followed by an overview of the work by Koolen et al. (2015). We then present suggestions for the use of such real-world analyses to make it easier to inject realism into SWTSs, followed by our discussion and suggestions for future work.

SIMULATED WORK TASK SITUATIONS

The use of assigned search tasks in IIR evaluation studies is a well-established testing protocol since the introduction of SWTSs (Borlund & Ingwersen, 1997). A SWTS is a short textual description that presents a realistic information-requiring situation that motivates study participants to search the IR system (Borlund, 2000; 2003). Borlund and Ingwersen (1997, pp. 227-228) describe how SWTSs can ensure realism by allowing for user interpretations of the task, leading to cognitively individual information need interpretations—as in real life. Furthermore, by being the same for all the study participants experimental control is provided, and the search interactions are comparable across the group of study participants for the same SWTS. Borlund (2003) formulates five requirements for ensuring realism in SWTSs: (1) tailoring the SWTS to the information environment and the participant group; (2) including participants' personal information needs; (3) randomizing the order of SWTSs to prevent biases; (4) pilot-testing the SWTSs; and (5) including the SWTSs when reporting on the study (Borlund, 2016, p. 396, 407). She further specifies that well-designed SWTSs are relatable and self-identifiable for participants due to contextual richness and overlap in topical interest (Borlund, 2016, p. 396). Such tailoring of SWTSs requires a certain degree of homogeneity of the group of study participants as well as a relatively high level of abstraction in the SWTS wording (Borlund, 2016, p. 397). Nevertheless, research shows that formulating realistic and engaging SWTSs remains challenging (Borlund, 2016).

RELEVANCE ASPECTS IN BOOK SEARCH

One way of injecting realism into SWTSs is by drawing inspiration from analyses of real-world information needs, such as the relevance aspects expressed by users, which determine what makes information items relevant to them. A more accurate understanding of real-world relevance aspects could help us formulate better SWTSs. In this paper, we draw inspiration from the

work by Koolen et al. (2015), who harvested 944 examples of complex book search requests from the LibraryThing forums to gain insight into the composition and complexity of the information needs represented by these requests. They analyze the relevance aspects expressed in these 944 book requests by adopting the seven categories identified by Reuter (2007) in her study of school children’s relevance aspects for book selection. Koolen et al. (2015, p. 188-189) add known-item search as the eighth relevance aspect. Table 1 shows an overview of the eight different aspects as well as their relative frequency in the 944 requests (third column). The right-most columns show how often the different relevance aspects co-occurred with each other. For example, the aspects *Familiarity* and *Engagement* co-occurred in 91 different requests.

Aspect	Description	% of requests (N = 944)	Co-occurrence frequencies							
			A	C	E	F	K	M	N	S
Accessibility	The language, length, or level of difficulty of a book.	16%	152	109	44	50	15	39	8	27
Content	Topic, plot, genre, style, or comprehensiveness of a book.	74%	698	172	267	100	176	26	99	
Engagement	Books that fit a certain mood, are considered high quality, or provide a certain reading experience.	23%	213	91	17	50	11	24		
Familiarity	Books similar to known books or related to a previous experience.	36%	338	12	83	17	45			
Known-item	Descriptions of known books with the sole purpose of identifying its title and/or author.	21%	202	85	0	1				
Metadata	Books with a certain title, by a certain author/publisher, in a particular format, or from a certain year.	28%	264	11	26					
Novelty	Books that are unusual or quirky, or have novel content.	4%	34	10						
Socio-cultural	Books related to the user’s socio-cultural background or values, have had a particular cultural or social impact or are popular or obscure.	14%	143							

Table 1. Relevance aspects in the book domain (Koolen et al., 2015), along with their relative occurrence in the 944 book requests and co-occurrence counts with other relevance aspects.

TAILORING SIMULATED WORK TASK SITUATIONS TO THE BOOK DOMAIN

A better understanding of the types of information needs in a particular domain and the relevance aspects expressed therein could help in the formulation of more realistic SWTSSs. While the most common aspects for book discovery—*Content* and *Familiarity*—may not necessarily be surprising relevance aspects to include in SWTSSs, other aspects such as *Engagement* and *Socio-cultural* are more of a ‘treasure’ in the view of the authors. They are not obvious elements to include, but can be very motivational to participants in that they involve feelings and values and may therefore be another dimension of interest. For example, users may search for funny books or books that challenge existing views (Koolen et al., 2015, p. 189). Knowledge of how to realistically combine different aspects—through co-occurrence frequencies—could also help with design of SWTSSs. We believe that knowledge of the relevance aspects in a domain could result in more realistic and better tailored SWTSSs.

DISCUSSION & CONCLUSIONS

In this poster paper, we have suggested how analyses of real-world information needs in specific domains can help inform the design of SWTSSs, more specifically in the book domain. Taking different relevance aspects and their relative frequency into account, as well as which aspects are commonly combined could provide inspiration for more realistic SWTSSs. For future work, we would like use the results of similar analysis of complex search requests in other domains to tailor SWTSSs and ultimately evaluate the quality of SWTSSs designed using inspiration from these different domains.

REFERENCES

Bogers, T., Gäde, M., Koolen, M., Petras, V., and Skov, M. (2018). "What was this movie about this chick?": A Comparative Study of Relevance Aspects in Book and Movie Discovery. *Proceedings of the 2018 iConference* (pp. 323-334).

Borlund, P. & Ingwersen, P. (1997). The Development of a Method for the Evaluation of Interactive Information Retrieval Systems. *Journal of Documentation*, 53(3), 225-250.

Borlund, P. (2000). Experimental Components for the Evaluation of Interactive Information Retrieval Systems. *Journal of Documentation*, 56(1), 71-90.

Borlund, P. (2003). The IIR Evaluation Model: A Framework for Evaluation of Interactive Information Retrieval Systems. *Information Research*, 8(3).

Borlund, P. (2016). A Study of the Use of Simulated Work Task Situations in Interactive Information Retrieval Evaluations: A Meta-Evaluation. *Journal of Documentation*, 72(3), 394-413.

Koolen, M., Bogers, T., Bosch, A. V., & Kamps, J. (2015). Looking for Books in Social Media: An Analysis of Complex Search Requests. *Proceedings of ECIR 2015* (pp. 184-196).

Reuter, K. (2007). Assessing Aesthetic Relevance: Children’s Book Selection in a Digital Library. *JASIST*, 58(12), 1745-1763.