The Electronic Library

Usability evaluation with tasks characterized by the information search process: The China National Knowledge Infrastructure

Cuichang Ma Shujin Cao Tinghua Gu

Received 13 May 2015
Revised 17 June 2015
Accepted 30 July 2015

Usability evaluation with tasks characterized by the information search process

The China National Knowledge Infrastructure

Cuichang Ma
Library, Sun Yat-sen University, Guangzhou, China, and
Shujin Cao and Tinghua Gu
School of Information Management, Sun Yat-sen University, Guangzhou, China

Abstract

Purpose – The purpose of this study is to present an approach to evaluating the usability of digital libraries in terms of knowledge states within the information search process (ISP) by investigating the connections among components with which users interact, perceived usability and task stages. The China National Knowledge Infrastructure (CNKI), the most popular digital library in China, was chosen to be studied.

Design/methodology/approach – With the research framework, measurements were generated for perceived usability in task stages for efficiency, effectiveness and satisfaction. Two usability experiments and usability tests were administered to 22 subjects in two groups. Three information retrieval tasks were designed according to the characteristics of knowledge states within the ISP for each experiment. The transaction logs were captured and the experimental results were recorded. The users completed a pre-search questionnaire and post-search questionnaire, and interviews were conducted.

Findings – The study showed that both interactions on components of the digital library and usability scores differed with sequential tasks characterized by knowledge states within the ISP; new usability problems relating to the CNKI were found based on the empirical results.

Originality/value – Many studies on the usability evaluation of digital libraries focused on the outcome rather than on the ISP. This study proposed an effective approach to usability evaluation based on users' knowledge states within the ISP.

Keywords Evaluation, Digital libraries, Information seeking behaviours, Usability engineering

Paper type Research paper

1. Introduction

To date, much of the research evaluating the usability of digital libraries (Jeng, 2006; Joo and Lee, 2011) has focused on the development of frameworks that describe, model and evaluate the users' interaction with digital libraries as a whole, across a set of measures.

This paper is supported by the National Social Science Fund of China (Grant No. 14CTQ015 and Grant No. 12&ZD222) in the entire research process. The authors are indebted to Associate Professor Yuxiang Zhao from Nanjing University and Penghui Lu from Wuhan University for academic support.

The Electronic Library
Vol. 34 No. 4, 2016
pp. 572-587

0264-0473
DOI 10.1108/EL-05-2015-0074

Such frameworks, however, have focused on the outcome of digital libraries as a whole rather than on the interaction process, and have not considered the components of digital libraries; in so doing, little attention has been given to the value of bottom-up evaluation (Balatsoukas *et al.*, 2009).

Some researchers have attempted to synthesize theories concerning both the information search process (ISP) and human–computer interaction to conceptualize the usability of digital libraries, such as Ferreira and Pithan (2005) and Makri *et al.* (2008). Interaction studies in the context of ISP, however, are always longitudinal because of the time span of the ISP, and so are inefficient at evaluating usability in terms of system design or improvement (Fuhr *et al.*, 2007).

Characterizing usability by how a digital library supports the interactions involved in the ISP, this paper aims to present an approach to evaluating the usability of the components of digital libraries, with tasks characterized by knowledge states in the ISP. Here, components are classified according to the interaction objects, including the navigation system, information organization system, search system and general components of the digital library as a whole. The following research questions were explored:

RQ1. Does the distribution of interactions on components of the digital library differ with the tasks characterized by the knowledge states in the ISP?

RQ2. Does the distribution of perceived usability differ among the tasks characterized by the knowledge states in the ISP?

RQ3. What are the usability problems with the research system?

With this aim, the China National Knowledge Infrastructure (CNKI) (www.cnki.net) was chosen as the focus of this study. It is the most popular digital library in China, encompassing a wide range of disciplines with a variety of genre types. Also, usability studies on CNKI are scarce in publications in China as well as worldwide, except for studies on user-centred design-related issues (Qiu, 2011), which would be comparable to the present study.

2. Literature review

2.1 Usability evaluation in the context of the information search process

Information-seeking theories, such as Kuhlthau's process model, Ellis's feature set and Dervin and Nilan's (1986) sense-making approach, could be integrated with information retrieval in an information storage and retrieval model, as Ingwersen and Järvelin (2005) proposed. On the operational level, Vakkari (2001) refined Kuhlthau's model of the information search process in the field of information retrieval (IR) and demonstrated that the information sought, search tactics, term choices and relevance judgements, as well as the contributing types of information in documents, depend systematically on the stage of the task performance process and the mental model of the searcher.

In terms of usability evaluation, Makri *et al.* (2008) even proposed the information behaviour (IB) method as an approach to evaluating functionality and usability of a digital library based on Ellis' information behaviour theory. Several other types of information behaviour research have been used to explore the relationship between the user's ISP and the digital library itself – or components of it – for usability evaluation. For example, Kim (2002) developed the information-seeking process model as a

framework for classifying usability problems. Ferreira and Pithan (2005) attempted to synthesize Kuhlthau's ISP theory with human–computer interaction for evaluating the usability of digital libraries. Based on the episode model (Belkin *et al.*, 1995; Cool and Belkin, 2002) and a four-phase interaction framework (Shneiderman *et al.*, 1997), Balatsoukas *et al.* (2009) proposed an evaluation framework for metadata surrogates in the process of relevance judgement.

Kuhlthau's ISP model is similar to Ellis's information behaviour model except that the ISP model reveals the changes in the user's cognitive and affective states, while Ellis's model directly links the user's behaviour to the components of the information system (Wilson, 1999). Thus, Ellis's information behaviour theory – linking behaviour and ISP – may be used to develop Kuhlthau's ISP theory.

2.2 Task stage classification and knowledge states

Because theories relating to the ISP may provide a theoretical means of evaluating usability, the classification of task stages is crucial. According to Kuhlthau (1988, 2004), the ISP is defined as the user's constructive activity of finding meaning from information to extend his or her state of knowledge, including initiation, selection, exploration, formulation, collection and presentation. Vakkari (2001) condensed the task performance stages proposed by Kuhlthau into three stages of IR, including before and after the focus formulation, as "methodological choice", for it was improbable that the authors would be able to differentiate all the phases in the four-month period of the proposal writing (Byström and Järvelin, 1995). Similar to Vakkari (2001), Byström and Hansen (2005) proposed that any task performance process can be divided into three stages: construction stage, performance stage and completing stage. Stages such as initial, selection and exploration incorporate the most difficult and time-consuming information activities – regarded as a weak state for the users (Palmer *et al.*, 2007); the existing ISP models that characterize the information-seeking behaviour in traditional fact-retrieval processes (Fu and Pirolli, 2007) need to be further developed.

In terms of knowledge states related to task stages, Kuhlthau (2004) stated:

> […] within the search process, thoughts evolved from unclear, vague uncertainty to clearer, more focused understanding or in Belkin's (1980) terms, from an anomalous state of knowledge to specificity.

Vakkari (2001) proposed that a "user's uncertainty decreases in their constructive activity of finding meaning from information to extend his or her state of knowledge on a particular problem or topic", as Kuhlthau (2004) also suggested. So, task stages could be defined as a change in users' knowledge states, which certainly occurs in the ISP model.

2.3 Usability constructs

Usability is a multidimensional construct and can be evaluated using various criteria. The most commonly cited definition is the one from the International Organization for Standardization (ISO) (1998) and they defined *usability* as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use". Nielsen (1993) summarized the five usability attributes as learnability, efficiency, memorability, errors and satisfaction. Similarly, but more specifically, Jeng (2006) proposed a usability model for academic digital libraries using four operational usability criteria – effectiveness, efficiency,

satisfaction and learnability – which is the same as the measurement instrument developed by Joo and Lee (2011).

Usability studies examine not only the surface representation of a system but also its underlying structures (Fuhr *et al.*, 2007) or components (Vrana, 2007; Xie, 2008). The information architecture theory (Rosenfeld and Morville, 2002), defined as the art and science of organizing and labelling information systems to support usability and findability (IAI, 2013), is widely used in user-centred design that support a range of information tasks (Batley, 2007; Toms, 2002). For example, Zhao (2007) and Xiang (2007) conceptualized usability constructs of digital libraries based on the information architecture theory, considering information organization, labelling, navigation and information search as components to be evaluated. In terms of usability evaluation for components of information systems, Thomsett-Scott (2007) conducted a competitive usability study for online catalogues of three university libraries. Fraser and Gluck (2008) investigated the usability of geospatial metadata or space-time matters.

Contextual factors, such as Web and system experience, and task knowledge have always been conceptualized in the evaluation of digital libraries in both information science and human–computer interaction (Nielsen, 1993). The influences of education level and academic background on information behaviour have been verified (Case, 2002). Tasks have been explored within the context of information searches or retrieval in studies such as those conducted by Ferreira and Pithan (2005) and Vakkari (2001).

2.4 Research framework

Based on a review of the above literature, this paper proposes a research framework as represented in Figure 1. The framework lists the main variables that would affect users' perception of usability by investigating the connections among components of the objects with which users interact, perceived usability and task stages. Information search tasks are defined according to the knowledge states in the ISP as controlled variables for investigating the proposed approach to usability evaluation.

3. Method

3.1 Variables and measures

While the ISP theories described the change in the users' states of thoughts, feelings, actions, uncertainty or states of knowledge in Belkin's (1980) terms during the ISP, they did not provide any direct specifications for usability design or evaluation. Therefore, there was still a gap between the stages of the information-seeking/problem-solving process that the models describe and the interactions the users performed at the system level for usability evaluation. To bridge this gap, the authors used the ISP models as lenses for mapping the information search tasks with characteristics of thoughts, uncertainty or knowledge states – the term used in this study – of the ISP as contextual factors in IR and for mapping the components of interaction objects with task stages. The variables of the research framework, operational definition and their measures are described below.

3.2 Task stages characterized by knowledge states in the information search process

Because Kuhlthau (2004) described the change of thoughts during the ISP as from ambiguity to specificity without clarifying them for each stage in a measurable manner, it was improbable to differentiate the change of the knowledge states within all the stages on an operational level. Based on Vakkari's (2001) study, the researchers chose to

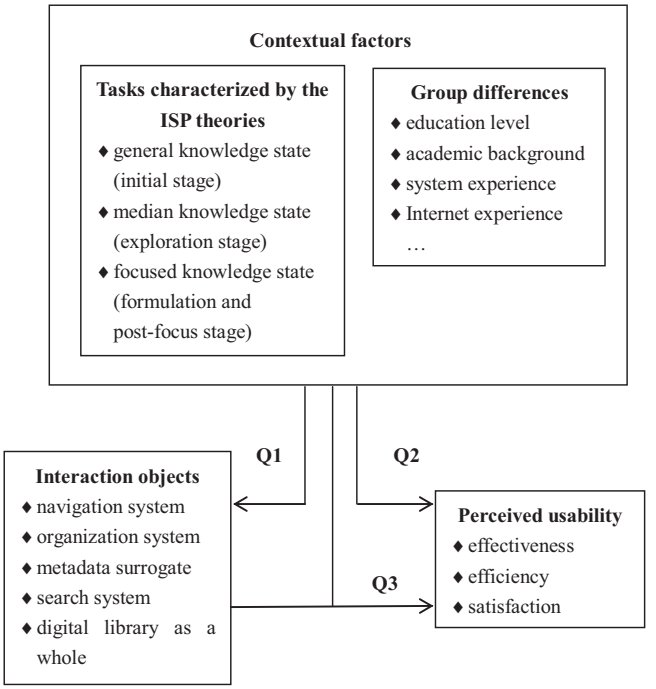


Figure 1.
Research framework of this study

condense the ISP into three stages according to the crucial points of the knowledge states – ambiguity, median and specificity – as the operational definition for task stages.

Thus, Vakkari's (2001) classification of task stages is revised as follows: the pre-focused stage is sub-divided into the initial and exploration stages, as the initial stage is difficult and time-consuming for a user in the knowledge state of ambiguity to begin the search with a general area, topic or problem; the focus formulation and post-focus stages are condensed into one stage to indicate a user's knowledge state of specificity when his/her focused perspective is formed; the exploration stage indicates a user's knowledge state of median – between ambiguity and specificity – to continue with the search so as to eliminate information that is inconsistent or incompatible with his/her inquiry. Thus, the six original stages in Kuhlthau's model are condensed into three, namely, the initial stage, the exploration stage and the focused formulation and post-focus stage, corresponding to the three sequential task stages, including general, median and focused knowledge states:

- The *general* state describes the knowledge state when a user begins the search with a general area, topic or problem (the "initial stage").
- The *median* state describes the knowledge state when a user continues with the search so as to eliminate information that is inconsistent or incompatible with his/her inquiry (the "exploration stage").
- The *focused* state describes the knowledge state when a user's focused perspective is formed (the "focused formulation and post-focus stage").

3.3 Interaction objects and the components with which users interact in task stages

Information architecture theory (Rosenfeld and Morville, 2002) may be constructed as the framework of interaction objects for investigating usability. A metadata surrogate system is regarded as distinct from an organization system as an independent interaction object to be studied. "Digital library as a whole" is used to indicate the components that have an overall effect on the entire process at any time. In the revised information architecture theory, definitions of interaction objects and their components to be studied are as follows:

- *Navigation system* means components designed to support browsing, including navigation of the entry pages, such as genre type, institution and publishing year, links to the reference and related articles and navigation of the result page, such as sorting and refinement of results (Rosenfeld and Morville, 2002; Xiang, 2007).
- *Organization system* means classification systems, including the disciplinary classification of entry pages, query pages and result pages according to the discipline (Rosenfeld and Morville, 2002; Xiang, 2007).
- *Metadata surrogate* means a substitute for an information item that contains data describing it from several aspects, including preview metadata surrogate and overview metadata surrogate (Balatsoukas *et al.*, 2009).
- *Search system* means components designed to support IR, including search fields and keyword lists (Rosenfeld and Morville, 2002; Xiang, 2007).
- *Digital library as a whole* means general components within other information objects that can cause usability problems at any time point in the entire process, including labelling systems, such as linkage, speed and tagging (Rosenfeld and Morville, 2002; Xiang, 2007).

With Ellis's information behaviour theory as a bridge (Ellis, 1989; Tibbo and Meho, 2003; Wilson, 1999), Table I shows interaction objects and their components mapped with task stages.

3.4 Usability constructs

Based on ISO 9241-11 (1998), *perceived usability* is defined as the extent to which components of digital libraries can support specified users to achieve goals with effectiveness, efficiency and satisfaction in tasks characterized by knowledge states in the ISP:

- *Effectiveness* means the ability to provide support effectively. Items of perceived effectiveness were rated on a usability scale according to whether users agreed with "the types of function are complete and the content is correct" for the components of interaction objects in sequential tasks.
- *Efficiency* means the ability to provide support in a timely and straightforward manner. Items of perceived efficiency were rated according to whether users agreed with "they are simple enough to understand" for the components of interaction objects in sequential tasks.
- *Satisfaction* means the degree of match between users' expectations and perception after usage. Items of perceived satisfaction were rated according to

Table I.
Summary of components of interaction objects in task stages

Task stages	Code	Interaction objects	Code	Components	Code	Reference
Initial	TI	Navigation system	N	Navigation of the entry page	N1	Kuhlthau, (2004);
				Links of reference and related articles	N2	Kuhlthau, (2008); Balatsoukas *et al.*, (2009); Tibbo and Meho (2003)
				Navigation of the result page	N3	
		Organization system	I	Disciplinary classification in entry page	I1	
		Search system	SE	Fields and keyword list	SE	
		Metadata surrogate	M	Overview metadata	M1	
Exploration	TE	Organization system	I	Disciplinary classification in query page	I2	Balatsoukas *et al.* (2009)
				Classification in the result according to discipline	I3	
		Navigation system	N	Navigation of the result page	N3	
		Metadata surrogate	M	Overview metadata surrogate	M1	
				Preview metadata surrogate	M2	
		Search system	SE	Fields and keyword list	SE	
Focus formulation and post-focus	TF	Metadata surrogate	M	Preview metadata surrogate	M2	Balatsoukas *et al.* (2009)
				Overview metadata surrogate	M1	
		Navigation system	N	Navigation of the result page	N3	
		Search system	SE	Fields and keyword list	SE	
At any time during the entire process	All stages	Digital library as a whole	G	Tagging in the interface	G1	Jeng (2006)
				Links	G2	
				Running speed	G3	

Note: Grey boxes indicate interaction objects or components added in the second phase

whether users agreed with "I am satisfied with them" for the components of interaction objects in sequential tasks.

- The overall satisfaction with the general usability of the interface and the system was measured, and users were asked to rate the labelling system of the interface, links and the running speed of the system.

Besides tasks, group differences are also examined as contextual variables in this study. The education level and academic background were factors controlled for in the subject recruitment of two groups, highlighting the differences between user groups in basic information, such as age, internet experience and system experience, and in the evaluation results, such as usability evaluation scores, span of each task and distribution of components in the ISP.

3.5 Experiment design and procedure

This study administered two controlled experiments and usability tests in two phases: Phase one was conducted in March 2013 in an information technology laboratory of the library at Sun Yat-sen University with six librarians; Phase two was conducted in May 2013 in a computer laboratory with 16 students from a school of computer science and education software at Guangzhou University. The reason for re-running the usability evaluation was to confirm the relationship among behaviours, interaction objects and usability discovered in Phase one in a larger sample with a variety of users. Librarians were recruited from the library of Sun Yat-sen University, and were given a gift worth RMB 30 (approximately equal to US$5) as compensation, while the students from Guangzhou University were required to finish the tasks as coursework for credit. It took about 1.5 h for a subject to finish the experiment, on average.

According to the operational definition of knowledge states within an ISP, tasks were designed based on the knowledge hierarchy of the "Chinese Classified Thesaurus" (Editorial Board of Chinese Classified Thesaurus, 2010) in the class of information description (G254.3) and in the class of software engineering (TP311.5). The order of the three tasks was fixed to simulate the change of topic scope within the ISP:

(1) Please introduce the research field of "information description (G254.3)"/"software engineering" (TP311.5).

(2) Please give the profile and hotspots of "automation of document indexing and cataloguing" (G254.36)/"software development" (TP311.52).

(3) Please give a review for the hotspot of "automation indexing" (G254.361)/"system analysis and design" (TP311.521).

The formal usability testing technique was used to evaluate the usability of the CNKI. An information pack was presented to each subject, including a tutorial manual, an announcement of research content, a pre-search questionnaire and three post-search questionnaires with a task list. Users were interviewed after the experiment for detailed information. A WebEx recorder was installed in the experimental computers to log the transactions while a camera was used to record the entire experiment for both phases.

The pre-search questionnaire was used to collect demographic data and information about the subject's familiarity with the system and internet experience. The post-search questionnaire was designed to test the subject's perception of usability after each task

and overall perception. The measure of usability was a five-point Likert scale, with 1 meaning "totally disagree" and 5 meaning "totally agree". According to the results in Phase one, the items of usability scales in the post-search questionnaires were revised in the Phase two for a better correspondence between the interaction objects, usability problems and task stages. Thus, the components for evaluation differed in both phases, though they shared the same theoretical framework.

4. Results

Usability problems were partly reported in the process of experiment and partly discovered from the results of evaluation. Usability problems and interactions based on content analysis of transaction logs were coded by two of the authors until complete agreement was reached. IBM SPSS Statistics 19.0 software was used for quantitative analysis of interaction and usability questionnaires.

4.1 Basic information and group differences

In total, 22 subjects were recruited in both phases. In Phase one, there were 6 subjects (4 female, 2 male), who were librarians with a master's degree; in the Phase two, there were 16 subjects (4 female, 12 male), who were university students studying computer science in grade two.

In terms of system experience, 50 per cent of the librarians used the CNKI once or twice per week; 50 per cent of the university students had never used the CNKI. In terms of search experience, 66 per cent of the librarians had over 10 years' experience, and the remainder had over 15 years' experience; 47.1 per cent of the university students had over 5 years' experience, 29.4 per cent had over 10 years' experience.

The results of the independent samples t-test showed that the system experience of the librarians was significantly higher than that of the university students ($t = -4.224$, df = 18, significance = 0.001); there was no significant difference between the two groups, though Web experience for most librarians was higher. The results of one-way analysis of variance on sex, system experience and search experience showed that these factors had no significant influence on the interaction time and perceived usability.

The search time for sequential tasks is shown in Figure 2. A descriptive analysis showed that the time for each subject for each group decreased with task stages, which resulted in a reduction in the average time. At each stage, the time duration for the librarians was greater than that for the university students (Figure 2). The interviews after the experiments showed that users in the first stage lacked domain knowledge, and thus had to acquire knowledge or cues from multi-angles and multi-ways, which was time-consuming.

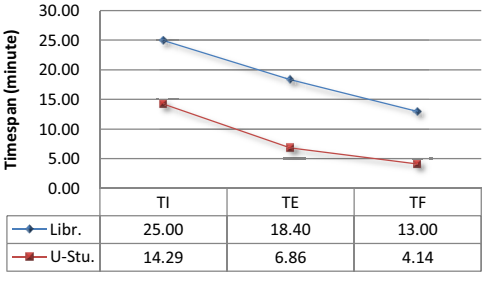


Figure 2.
Time for sequential task stages

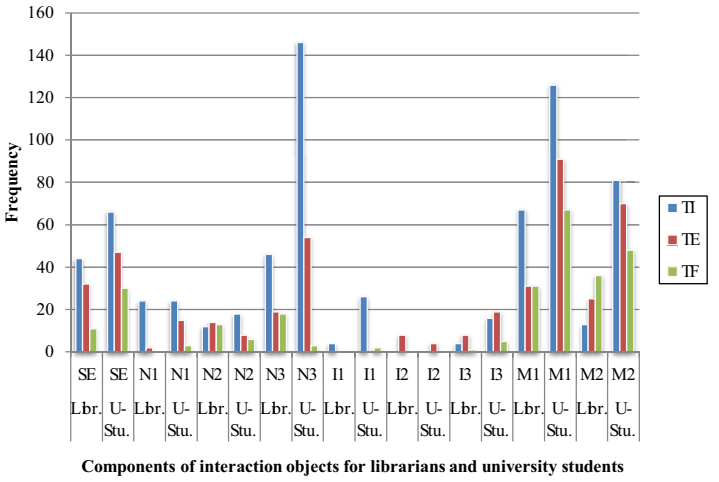
The results of independent samples *t*-tests showed that the average time duration for the librarians was significantly greater than that for the university students ($t = -5.971$, df = 18, significance = 0.000). Specifically, the time duration for the librarians in the TE ($t = -3.795$, df = 18, significance = 0.001) and TF ($t = -5.332$, df = 18, significance = 0.000) stages was significantly greater than that for the university students.

4.2 RQ1*: distribution of interactions on components in sequential tasks*

To answer *RQ1*, this study investigated the frequency and distribution of interaction times for the components of interaction objects in task stages.

The frequency of interaction times on the components in all tasks for both groups is shown in Figure 3. The interaction with navigation of the result pages (N3, 46 for librarians, 146 for university students) and overview metadata surrogates (M1, 67 for librarians, 126 for university students) showed the highest frequency in the initial stage (TI) for both groups; the overview metadata surrogates (M1) and preview metadata surrogates (M2) were high in all stages for both groups. Generally, the interaction with the organization system (including I1, I2 and I3) was low in all three stages for both groups.

The distribution of the components in sequential tasks is shown in Figure 4. The descriptive analysis revealed that interactions on information search (SE), entrance navigation (N1), navigation of the result page (N3) and overview metadata surrogate (M1) all decreased with sequential tasks; interactions on the links of reference and related articles (N2) were almost the same in all tasks for librarians, and decreased in the focus formulation and post-focus stage for university students; interactions on the preview metadata surrogates increased for the librarians, but decreased for the university students; and interactions on entrance classification (I1) and classification for search (I2) were mainly concentrated in the first and second task stages.



Note: Meanings of codes of components and task stages are indicated in Table I

Figure 3.
Frequency of interaction times on components in sequential tasks

In terms of preview metadata surrogates (M2) that contain the main information about literature, the interaction times for the librarians increased with the sequential tasks, but decreased for the university students. The university students were not as familiar with domain knowledge as the librarians with a master's degree, who would spend more time and interact more with the preview metadata surrogates to advance their learning in the initial stage. The librarians tended to focus on the content of literature in the exploration and focus stages with specific goals, leading to greater interaction time durations in these stages.

4.3 RQ2*: Distribution of perceived usability in sequential tasks*

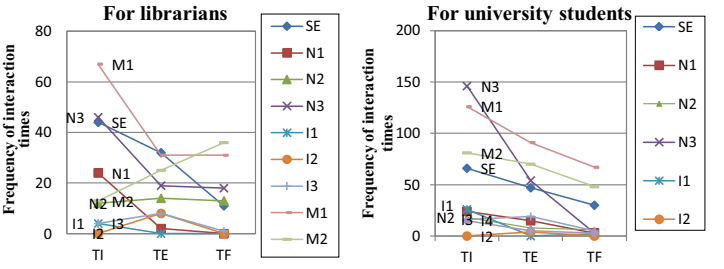
To answer *RQ2*, this study examined the distribution of scores of usability in sequential tasks for both groups and for components. The usability score for each component was multiplied by the effectiveness, efficiency and satisfaction of each subject. The usability score in each stage was the average score of components in the same stage for each subject, which could be calculated according to equation (1), where *C* is the construct of usability for components, including efficiency (C_1), effectiveness (C_2) and satisfaction (C_3), and *m* stands for the number of components in each task stage:

$$S_c = \frac{\sum_{n=1}^{m} C_{1n} \times C_{2n} \times C_{3n}}{m} \quad (1)$$

In terms of general usability, problems existed in all tasks of the evaluation, so they were considered as a whole. A general usability score for each subject was calculated according to equation (2), where S_G is the general perception of usability, including tagging (G1), links (G2) and speed (G3):

$$S_G = G_1 \times G_2 \times G_3 \quad (2)$$

The average score of the three stages and general usability was regarded as the total scores of the CNKI for each subject, which could be calculated according to equation (3), where *j* is the number of tasks:



Note: Meanings of codes of components and task stages are indicated in Table I

Figure 4.
Distribution of components for librarians and university students in sequential tasks

$$S = \left(\sum_{i=1}^{j} \frac{\sum_{n=1}^{m} C_{1n} \times C_{2n} \times C_{3n}}{m} + G_1 \times G_2 \times G_3 \right) / j + 1 \quad (3)$$

The usability scores of both groups are listed in Table II. The distributions of the usability scores in sequential tasks are listed in Figure 5. Where the items of scales of usability differed between the two groups, the total scores of usability were not comparable. Nevertheless, the usability scores were found to decrease with sequential tasks for both groups, indicating that the users' perceived usability increased with increasing domain knowledge and the specialization of the task. This result agrees with the aforementioned distribution of interactions on components, indicating that perceived usability differs within the ISP. Also, the general usability scores were lower than the average.

Furthermore, the average scores were higher for the librarians than for the university students. The distribution of the average scores agrees with the aforementioned distribution of time durations, revealing the difference in perception of usability between the two groups.

At the component level, the scores for navigation of the entry page (M = 37, SD = 20.85) and disciplinary classification of the entry page (M = 22.33, SD = 9.42) in the initial stage, disciplinary classification of the query page (M = 37.5, SD = 12.83) and disciplinary classification of the result page (M = 27.33, SD = 18.95) in the exploration stage and navigation of the result page (M = 38, SD = 19.88) in the focus formulation and post-focus stage were all lower than average for the librarians. The scores for navigation of the entry page (M = 15.69, SD = 16.54), search fields and word lists (M = 13, SD = 9.37), overview metadata surrogates (M = 12, SD = 7.37) and navigation of the result page (M = 12.88, SD = 8.62) in the initial stage and for search fields and word lists (M = 27.33, S = 18.95) in the exploration stage were lower than average for the university students. Thus, both groups shared a common low rating for the navigation of the entry page.

Group	Usability	TI	TE	TF	G	Average score
Librarians	Mean	37.6	40.8	49.8	36.00	41.1
	SD	11.1	12.5	10.6	13.2	6.2
Students	Mean	12.7	18.0	21.3	16.4	17.1
	SD	10.5	23.4	10.9	8.7	3.6

Table II. Scores of usability for librarians and university students

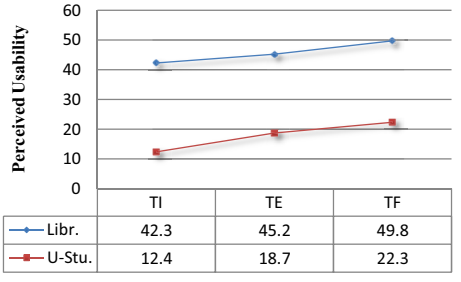


Figure 5. Distribution of usability scores for both groups in the ISP

4.4 RQ3*: Usability problems with the China National Knowledge Infrastructure*

Usability problems are summarized according to the interaction objects within the ISP with suggestions for improvement as follows: some of these have not been previously reported, especially from the ISP perspective.

The frequency of interaction times for the organization system (including I1, I2 and I3) was low in all stages for both groups. Nevertheless, disciplinary classification of the entry page as part of the organization system was rarely used by the librarians and never used by the university students. The interviews after the experiments found that librarians were familiar with disciplinary classifications, and were willing to use them. However, the general classification presented on the entry page was reported to hinder use, as it was too simple and needed clicks to present a second classification, while the users were uncertain as to what information they could get from it. Thus, it would be better if the CNKI presented the disciplinary classification on the entry page with a second- or third-level classification explicitly to facilitate users in the initial and exploration stages to find what they wanted from a general topic.

In terms of the navigation system, the functional link for sorting on the result page was not embedded in the headings of the table for the overview metadata surrogates, which led to unproductive clicks in the exploration and focus formulation stages. It would be better if the CNKI added functional links to the headings of the table to enable the users to sort as they see it.

The metadata system was a highly rated interaction object in all task stages, which warrants further explanation. Though the scores for the metadata system were higher than average in most stages, some usability problems were discovered, which have also been reported by Qiu (2011). The preview metadata surrogates were regarded as too simple for relevance judgement; reference to the relevant article in the preview metadata surrogates covers such a lengthy page that it is better to open them in a new page. Thus, it would facilitate the users if the CNKI redesigned the architecture of the preview metadata surrogates to support relevance judgement and added links for the relevant article to save time scrolling up and down a long page. This is important for the users who are in a weak knowledge state and even anxious in the initial and exploration stages.

The usability scores showed that the perception of usability improves with sequential tasks, highlighting that it is better to focus on usability evaluation and improvement of components in the initial and exploration stages. The group differences in the usability scores indicated that the librarians had a higher satisfaction with usability than the university students, for the CNKI experience was rated significantly higher by the librarians than by the university students. Thus, the usability problems of new users merit further attention. The scores for navigation of the entry pages in the initial stage were lower than average for both groups. Thus, it would be preferable to re-arrange components, such as genre type, institution, publishing year or sources, according to the interaction and perceived usability of the ISP to improve the usability of the system.

5. Conclusion

As previous usability studies focused on the outcome of a digital library as a whole, rather than the interaction process with its components, this study presents an alternative approach to the usability evaluation of digital libraries based on the theories

about the ISP. A research framework was constructed with the empirical results, highlighting the differences in interaction and usability across three knowledge states within the ISP. Specifically, the frequency of interaction times decreased with the sequential tasks for most components; the usability scores increased, while the time durations decreased with the sequential tasks; new usability problems with the CNKI were found based on the empirical results.

Compared with the studies relating to the task stages as longitudinal studies, this study takes further steps to explore the relationship between the task stages and usability evaluation, attempting to divide the work tasks within the ISP into three sub-tasks characterized by knowledge states to improve the efficiency of process-based usability evaluation in one experiment. Our results agreed with the characteristics of the ISP and Palmer *et al.*'s (2009) weak information states. Nevertheless, other dimensions of the ISP relating to usability evaluation need to be further explored.

The implications for usability design or improvement are derived from the findings that it are necessary to conduct usability evaluation according to the group differences and interaction within the ISP for digital libraries. The group differences, however, are just one way of verifying the proposed approach; contingency factors, such as differences in digital libraries, users and conditions of experiments, must also be considered. Therefore, research on a larger scale that takes into account more factors is necessary in the future.

References

Balatsoukas, P., Morris, A. and O'Brien, A. (2009), "An evaluation framework of user interaction with metadata surrogates", *Journal of Information Science*, Vol. 35 No. 3, pp. 321-339.

Batley, S. (2007), *Information Architecture for Information Professionals*, Chandos Publishing, Oxford.

Belkin, N.J. (1980), "Anomalous states of knowledge as a basis for information retrieval", *Canadian Journal of Information Science*, Vol. 5, pp. 133-143, available at: https://smallbusinessownerib.files.wordpress.com/2011/11/belkin1980.pdf

Belkin, N.J., Cool, C., Stein, A. and Thiel, U. (1995), "Cases, scripts and information-seeking strategies: on the design of interactive information retrieval systems", *Expert Systems with Applications*, Vol. 9 No. 3, pp. 379-395.

Byström, K. and Hansen, P. (2005), "Conceptual framework for tasks in information studies", *Journal of the American Society for Information Science and Technology*, Vol. 56 No. 810, pp. 1050-1061.

Byström, K. and Järvelin, K. (1995), "Task complexity affects information seeking and use", *Information Processing and Management*, Vol. 31 No. 2, pp. 191-213.

Case, D.O. (2002), *Looking for Information: A Survey of Research on Information Seeking, Needs and Behavior*, Academic Press, San Diego, CA.

Cool, C. and Belkin, N.J. (2002), "A classification of interactions with information", *Proceedings of the Fourth International Conference on Conceptions of Library and Information Science, Libraries Unlimited, Greenwood Village, CO.*

Dervin, B. and Nilan, M. (1986), "Information needs and uses", *Annual Review of Information Science and Technology*, Vol. 21, pp. 3-33.

Editorial Board of Chinese Classified Thesaurus (2010), *Chinese Classified Thesaurus*, available at: http://cct.nlc.gov.cn/login.aspx (accessed 2 January 2013).

Ellis, D. (1989), "A behavioural approach to information retrieval system design", *Journal of Documentation*, Vol. 45 No. 3, pp. 171-212.

Ferreira, S.M. and Pithan, D.N. (2005), "Usability of digital libraries: a study based on the areas of information science and human-computer interaction", *OCLC Systems & Services*, Vol. 21 No. 4, pp. 311-323.

Fraser, B. and Gluck, M. (2008), "Usability of geospatial metadata or space-time matters", *Bulletin of the American Society for Information Science*, Vol. 25 No. 6, available at: www.asis.org/Bulletin/Aug-99/fraser_gluck.html (accessed 4 March 2014).

Fu, W.T. and Pirolli, P. (2007), "SNIF-ACT: a cognitive model of user navigation on the World Wide Web", *Human-Computer Interaction*, Vol. 22 No. 4, pp. 355-412.

Fuhr, N., Tsakonas, G., Aalberg, T., Agosti, M., Hansen, P., Kapidakis, S., Klas, C.P., Kovacs, L., Landoni, M., Micsik, A., Papatheodorou, C., Peters, C. and Solvberg, I. (2007), "Evaluation of digital libraries", *International Journal on Digital Libraries*, Vol. 8 No. 1, pp. 21-38.

Information Architecture Institute (IAI) (2013), *What is Information Architecture?*, available at: www.iainstitute.org/documents/learn/What_is_IA.pdf (accessed 20 February 2014).

Ingwersen, P. and Järvelin, K. (2005), *The Turn: Integration of Information Seeking and Retrieval in Context*, Springer, Netherlands.

International Organization for Standardization (ISO) (1998), *ISO 9241-11: Ergonomic Requirements for Office Work with Visual Display Terminals: Part 11: Guidance on Usability*, International Organization for Standardization, Geneva.

Jeng, J.H. (2006), "Usability of the digital library: an evaluation model", PhD thesis, State University of New Jersey, Rutgers, NJ.

Joo, S. and Lee, J.Y. (2011), "Measuring the usability of academic digital libraries", *The Electronic Library*, Vol. 20 No. 4, pp. 523-537.

Kim, K. (2002), "A model of digital library information seeking process (DLISP model) as a frame for classifying usability problems", PhD thesis, State University of New Jersey, Rutgers, NJ.

Kuhlthau, C.C. (1988), "Developing a model of the library search process: cognitive and affective aspects", *Reference Quarterly*, Vol. 28 No. 2, pp. 232-242.

Kuhlthau, C.C. (2004), *Seeking Meaning: A Process Approach to Library and Information Services*, Libraries Unlimited, Westport, CT.

Kuhlthau, C.C., Heinström, J. and Todd, R. (2008), "The 'information search process' revisited: is the model still useful?", *IR Information Research*, Vol. 13 No. 4, available at: www.informationr.net/ir/13-4/paper355.html (accessed 2 February 2015).

Makri, S., Blandford, A. and Cox, A. (2008), "Using information behaviors to evaluate the functionality and usability of electronic resources: from Ellis's model to evaluation", *Journal of the American Society for Information Science and Technology*, Vol. 59 No. 14, pp. 2244-2267.

Nielsen, J. (1993), *Usability Engineering*, Academic Press, Boston, MA.

Palmer, C.L., Cragin, M.H. and Hogan, T.P. (2007), "Weak information work in scientific discovery", *Information Processing & Management*, Vol. 43 No. 3, pp. 808-820.

Palmer, C.L., Teffeau, L.C. and Pirmann, C.M. (2009), *Scholarly Information Practices in the Online Environment: Themes from the Literature and Implications for Library Service Development*, OCLC Research and Programs, Dublin, OH.

Qiu, M.H. (2011), "Research on digital library design based on user experience", *Journal of Intelligence*, Vol. 30 No. 7, pp. 163-168.

Rosenfeld, L. and Morville, P. (2002), *Information Architecture for the World Wide Web*, 2nd ed., O'Reilly & Associates, Sebastopol, CA.

Shneiderman, B., Byrd, D. and Croft, B. (1997), "Clarifying search: a user-interface framework for text searches", *D-Lib Magazine*, Vol. 3 No. 1, pp. 18-20.

Thomsett-Scott, B. (2007), "The thrills and chills of employing competitive usability in a study of online public access catalogs", *Technical Services Quarterly*, Vol. 25 No. 1, pp. 27-37.

Tibbo, L. and Meho, H. (2003), "Modeling the information-seeking behavior of social scientists: Ellis's study revisited", *Journal of the American Society for Information Science and Technology*, Vol. 54 No. 6, pp. 570-587.

Toms, E.G. (2002), "Information interaction: providing a framework for information architecture", *Journal of the American Society for Information Science and Technology*, Vol. 53 No. 10, pp. 855-862.

Vakkari, P. (2001), "A theory of the task-based information retrieval process: a summary and generalization of a longitudinal study", *Journal of Documentation*, Vol. 57 No. 1, pp. 44-60.

Vrana, R. (2007), "The importance of usability in development of digital libraries", INFuture2007: Digital Information and Heritage, available at: http://infoz.ffzg.hr/INFuture/2007/pdf/5-01%20Vrana,%20Radovan,%20The%20Importance%20of%20Usability%20in%20Development%20of%20DL.pdf (accessed 10 April 2015).

Wilson, T.D. (1999), "Models in information behaviour research", *Journal of Documentation*, Vol. 55 No. 3, pp. 249-270.

Xiang, J.J. (2007), "Study on usability evaluation of University Digital Library in China based on information architecture rules", PhD thesis, Xinan University, Chongqing.

Xie, H.I. (2008), "Users' evaluation of digital libraries (DLs): their uses, their criteria, and their assessment", *Information Processing and Management*, Vol. 44 No. 3, pp. 1346-1373.

Zhao, Y.X. (2007), "Study of evaluation on information architecture usability for public library website", *New Technology of Library and Information Service*, Vol. 2 No. 3, pp. 60-64.

Further reading

Ellis, D., Cox, D. and Hall, K. (1993), "A comparison of the information seeking patterns of researchers in the physical and social sciences", *Journal of Documentation*, Vol. 49 No. 4, pp. 356-369.

Ellis, D. and Haugan, M. (1997), "Modelling the information seeking patterns of engineers and research scientists in an industrial environment", *Journal of Documentation*, Vol. 53 No. 4, pp. 384-403.

About the authors

Cuichang Ma is currently an Associate Professor at the library of Sun Yat-sen University in Guangzhou, China. She is pursuing a doctoral degree, and has a master's degree from the School of Information Management at Sun Yat-sen University. Her research interests include usability evaluation and user–system evaluation.

Shujin Cao is currently a Professor and the Dean of the School of Information Management at Sun Yat-sen University. His research interests include information retrieval and user behaviour. Shujin Cao is the corresponding author and can be contacted at: caosj@mail.sysu.edu.cn

Tinghua Gu is currently a Doctoral Student in the School of Information Management at Sun Yat-sen University. Her research interests include user behaviour and knowledge organization.