Learning outcomes during information search in digital archives

I-Chin Wu
Graduate Institute of Library and Information Studies, School of Learning Informatics, National Taiwan Normal University, Taiwan
icwu@ntnu.edu.tw

Bo-Xian Huang
Graduate Institute of Library and Information Studies, School of Learning Informatics, National Taiwan Normal University, Taiwan
bxhuang33@gmail.com

Pertti Vakkari
Faculty of Information and Communication Technology, Tampere University, Finland
pertti.vakkari@tuni.fi

ABSTRACT

A museum's digital archive system gathers information about cultural heritage and makes it accessible to the public. In this study we clarify the extent to which search behaviors reflect task outcome and foster users' knowledge of painting and calligraphy. Ten users participated who are special interest museum visitors. They joined in this evaluation of the Digital Archives of Chinese Painting and Calligraphy Search System (DA-PCSS) of the National Palace Museum in Taiwan. Participants' search activities and interactions with the DA-PCSS were recorded in two simulated tasks. The results show that the high-performance (HP) group who received high scores for their essays on the tasks formulated precise and relevant queries. Furthermore, the HP group were able to find information needed inside or outside the system to explore and synthesize the information, which was clearly reflected in their search move patterns, during the search process. Our results suggest that the criteria for learning at various stages of search suggested by Vakkari (2016) seem to validly reflect the quality of the search outcomes. In all, the results elucidate how the evaluated system supports users as they search for target items, as well as how learning occurs during the search process and in turn influences task outcomes.

KEYWORDS

Chinese painting and calligraphy; Digital archives search system; Information search process; Meaning making process; Search as learning

INTRODUCTION

Using information technology, museums can present their collections digitally and extend the benefits of physical museums to the public. These museums as culture heritage institutions attract wider and more diverse user groups, especially for general public and non-professional users (Skov, 2013; Walsh, Hall, Clough, & Foster, 2020). Following the trend of open digital collections in museums and libraries, the National Palace Museum (NPM) (https://www.npm.gov.tw/)—the most important and popular museum in Taiwan—established database searches in 2015. According to its annual report, as of December 31, 2020, the NPM had a total of 698,796 works in its collection. Among these, 80% of the works in the NPM's Chinese Painting and Calligraphy collection have been digitalized and put online for public use (https://painting.npm.gov.tw/). The NPM has not only cultivated excellent digital archive resources on the Internet but has also begun to display Taiwan's rich cultural assets and make them both searchable and available to the public (Chen, Hsang, Chiang, & Hong, 2002). Digitization of museum collections allows users to easily and accurately find information for further research and use, and has given the public access to data previously limited to museum professionals. However, even if museums digitize their collections and make them public, can users effectively access the data therein, and are they aware of what content is available in the resources? We seek to explore the users' museum online experience by examining learning outcomes during information search.

Skov and Ingwersen (2014) investigate characteristics of the web search behavior among visitors to the National Museum of Military History in Copenhagen. The authors identified four characteristics of online museum visitors: a high propensity for visual experiences, exploratory search behavior, broad known item and/or element searches, and meaning-making. Exploratory search behaviors generally occurred while completing topic search tasks, and the visual aspect of search behavior was shown to directly support those behaviors. In addition, the participants used their background knowledge of military history to help them to interpret new clues and engage in further meaning-making behaviors. By extension, their study inspired the design of tasks in our research and our exploration of special interest museum visitors' search behaviors.

84th Annual Meeting of the Association for Information Science & Technology | Oct. 29 – Nov. 3, 2021 | Salt Lake City, UT.

To date, research in online museums has focused on evaluating the usability of the systems and users' characteristics and motivations for visiting digital museums (Falk, 2013; Pallas & Economides, 2008; Walsh et al, 2020). By contrast, few studies have investigated how knowledge formation occurs as a result of information searches (Rieh, Gwizdka, Freund, & Collins-Thompson, 2014). We examine whether users learn during the search process and display meaning-making behaviors based on learning criteria derived from information searches as in Vakkari's (2016) study. To that purpose, we explored users' search and interaction patterns in three search stages (search formulation, source selection, and source interaction) between groups with different task-performance outcomes, to reveal which characteristics promote learning outcomes. We adopt zero-order state transition matrices (ZOSTs) and lag sequential analysis (LSA) to calculate the frequency (i.e., probabilities) of search behaviors for users with different task outcomes (Bakeman, & Gottman, 1999; Sackett, 1979; Wildemuth, 2004). In this study, we seek to determine the characteristics of users' search and interaction patterns in the three stages of searching for information for users with different task performance, and to understand which promote better or worse learning outcomes as reflected in the final stage of the process (i.e., synthesizing and presenting information).

INFORMATION SEARCH BEHAVIOR FOR ONLINE MUSEUMS

The growth of online museums in the 1990s was international and exponential as evidenced in research evaluating different types of online museums and their visitors' motivations and demographics (Falk, 2013; Skov & Ingwersen, 2014; Walsh et al., 2016). Previous studies regard physical and online museums as complementary; however, they are characterized by different collection information environments that motivate visitors to visit them for various different reasons (Goldman & Schaller, 2004; Marty, 2007; Ross & Terras, 2011). Here, we review the related research to understand the information needs of visitors to various types of online museums by attending to their motivation or search behaviors.

Goldman and Schaller (2004) survey the relationship between visitors' motivations for visiting museum websites and their meaning-making processes. Teachers and students who were selected by this study visited a website to find specific or meaningful information and used four different types of museums. There were 1166 respondents in total to the survey. General visitors were excluded from the research target. The research results show that the intent to search for specific information (20.4% of respondents on average for the four museums) and to explore interesting information (17.3% of respondents) are the main reasons to visit an online museum instead of planning a trip to its physical location, as documented also in previous research (Sarraf, 1999).

Ross and Terras (2011) conduct an online survey and two simulated search tasks to analyze the information-seeking behaviors of scholars utilizing the British Museums' Collection Online (COL). Their work explores the value of online museums from the perspective of visitors' characteristics, their search strategies, and the usability of the system. A total of 2657 responses were received from a random sample of visitors from 57 countries. Among these, more than 50% of visitors identified their motivation for using the COL as undertaking academic research or exploring personal interests, which shows the scholarly value of visiting online museum digital resources instead of a physical visit, as mentioned in other studies. The known object type and free text searches are important functions for exploring online collections. Interestingly, regardless of whether the respondents left positive comments about the content of the system, they were unable to achieve good performance on the simulated search tasks. That is, although scholars use precise search terms, there is a gap in their understanding of the search terms and the return of the system's content. They suggest that a user-friendly metadata schema is required for further development of the system.

Early studies of online museums focus on the usage motivations and information needs of academic users of digital museum resources. However, online museums serving as culture heritage institutions attract wider and more diverse user groups. This applies especially to special-interest museum visitors categorized by Skov's (2013) study or the general public and non-professional users, as investigated by Walsh, Hall, Clough, and Foster's (2020) study. Skov (2013) uses triangulation as a qualitative research method for investigating visitors' hobby-related information-seeking behaviors in the Military Museum's collection database in Denmark. Web questionnaires collected from 132 respondents were analyzed to gain a better understanding of their leisure context areas and how they used digital museum resources in their daily lives. Twenty-four participants were classified into two groups of hobbyists—collectors and enthusiasts—as described by Stebbins's (2007) taxonomy of types of leisure pursuit of hobbies, and were invited to a follow-up interview. The research finds that most collectors have specific known-item needs, whereas most liberal arts enthusiasts have exploratory topical information needs. The contribution of this work is to find out what type of visitors cross the two hobby types, that is, who conduct more than just general exploration of a hobby or interest search. Accordingly, the research characterizes participants as special-interest museum visitors pursuing a long-standing interest or hobby, i.e., everyday information-seeking behaviors.

Walsh, Clough, and Foster (2016) review articles to categorize users of a digital cultural heritage (DCH) system into six groups from the generic perspectives of domain knowledge, technical expertise, and motivation. Their study reveals that learning, leisure, and planning visits are major motivations of DCH system users. For example, users with a high level of domain knowledge and a medium level of technical skills are similar to visitors at a special interest museum as characterized by Skov and Ingwersen (2014). The group of users who spent more time exploring the online museum as a personal hobby required more detailed information about the museum's collections. As such, the findings reveal that different types of users require different information or types of system support to fulfill their information-related needs.

Using survey or simulated search task methods, these studies provide insight into the information needs of different types of online museum visitors from the perspective of their motivation or search behaviors. However, there is still a paucity of empirical research into the issue of how learning occurs during the search process and in turn how it influences task outcomes for the general public or special-interest museum visitors. We seek to fill this research gap by conducting an empirical study based on learning criteria derived from information searches as in Vakkari's (2016) study via a simulated work task approach. We seek a better understanding of the extent to which the system supports visitors' learning while searching for information online.

METHODOLOGY

To observe the search processes, we used usability testing software to record participant's interactions with the Digital Archives of Chinese Painting and Calligraphy Search System (DA-PCSS). Two simulated tasks were designed based on two types of knowledge: tasks related to factual knowledge and conceptual knowledge respectively. Herein, we briefly explain the components of the evaluation framework in Figure 1.

Learning tasks and participants: We designed the two tasks based on Borlund (2003), as shown in Figure 2. Note that Task 1 concerns fact-based knowledge, for which the system provides less information support, whereas Task 2 concerns concept- and relationship-based information that can be retrieved from the system. We recruited ten users to participate in the DA-PCSS evaluation. We invited participants who were interested in painting and calligraphy to take the time to learn more about the paintings and calligraphy via the system. All of the participants had visited the museums at least twice during the preceding year and have been using the DA-PCSS system or NPM website to look for information. They also showed long-term personal interest in painting or calligraphy. After accomplishing the tasks, we also asked the participants to produce ratings for their perceived prior knowledge of each task (ranging from 0 to 100).

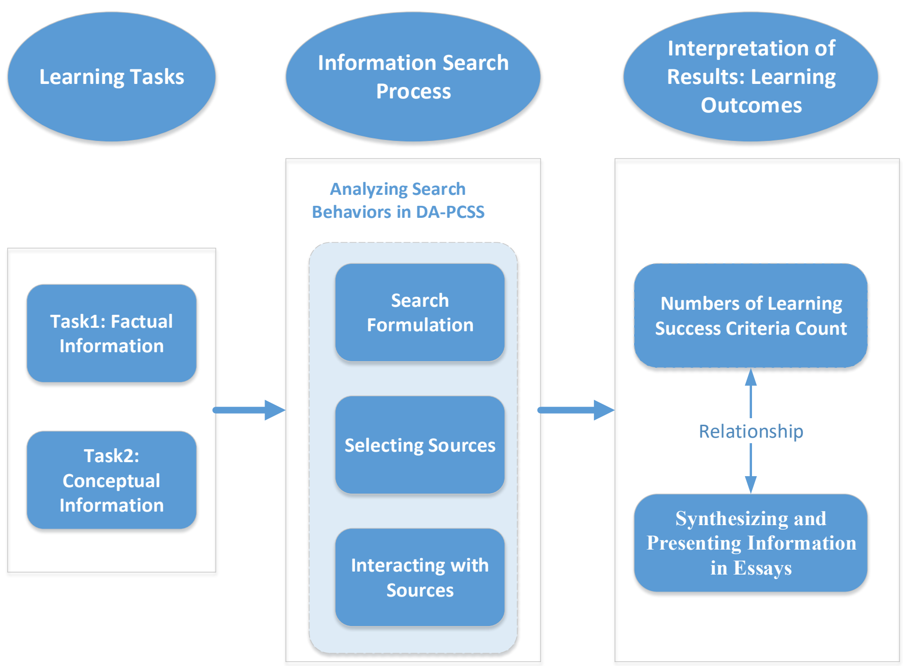

Figure 1. Evaluation Framework

Task Description

Task 1: While watching the flower painting exhibition, you see the "boneless" technique, which piques your interest in the characteristics of the boneless paintings and related works.

※Please record your learning outcome in writing: describe the characteristics of the "boneless" technique, for example, the author, works, topic and dynasties of the "boneless" style, the meaning of certain techniques, etc., or other content that you are interested in.

Task 2: As you browse the website, you find Fan Kuan's "Travellers among mountains and streams", which is one of the Three Treasures of the Palace Museum. The others are Guo Xi's "Early Spring" and Li Tang's "Wind in the Pines Among a Myriad Valleys". All three are landscape paintings, and you use the Antique digital archive retrieval system to learn more about these three works. The Three Treasures of the Palace Museum is shown in below.

※Please record the learning outcome in writing: similarities and differences, look for one to three paintings with the same techniques as the three paintings and the author, and other content you are interested in and describe the concepts you have learned.

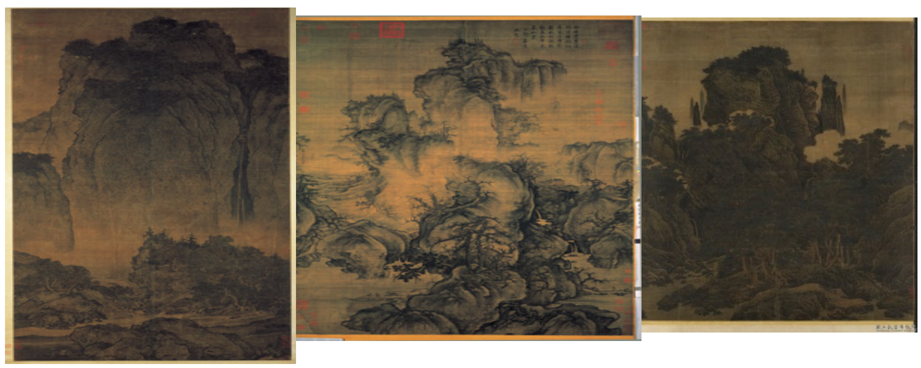

Figure 2. Task Description and Three Treasures in the Palace Museum (Source: https://painting.npm.gov.tw/)

Online search behavior analysis: We used Morae software to record participants online search behaviors. Each task lasted around 40 minutes on average. Their task performance and learning outcomes were not only reflected in their synthesis of presenting information in essays but also reflected in the number of achieved learning criteria per Vakkari's (2016) study. Vakkari (2016) has hypotetized how increasing learning, i.e. the growth of knowledge, across search sessions is reflected in users' search behavior including search formulation, selecting sources and interacting with sources for task outcome. He proposes that the hypothetized changes in search behavior can be used as indicators of learning across sessions. Although the hypotheses concern changes between search sessions, we apply them within search sessions for inferring criteria of learning. After analyzing participants' search behavior and learning outcomes, we differentiated them into better and worse task-performance groups. We then adopted zero-order state transitions (ZOSTs) to calculate the frequency of search paths in the target system (reference). For the ZOSTs, the frequency of each transition and its proportion in each set of transitions were calculated in the ZOSTs table. We labeled the ZOST results based on the codes shown in Table 1. For the ZOSTs, the frequency of each transition and its proportion in each set of transitions were calculated in the zero-order state transition table. Furthermore, we used lag sequential analysis (LSA) to derive the significant repeating patterns from search paths. The LSA method explores, summarizes, and statistically tests cross dependencies between behaviors that occur in interactive sequences (Sackett, 1979). LSA results facilitate an understanding of paths that are significant in terms of path frequency and path transfer probability, allowing us to analyze participants' task behaviors beyond descriptive statistic information.

Information search process and learning criteria: We analyzed participants' search and interaction patterns over the three search stages (search formulation, source selection, and source interaction) while accomplishing the two tasks. The observed items are listed in Table 1. Each participant was to write a short essay to present the task results for scoring by an expert. The expert's field of expertise was fine art and he/she had engaged in art-related work for over ten years. The expert evaluated the essay of each task primarily based on the number of relevant concepts/topic that were addressed and the related knowledge on each concept/topic that was explained clearly. In addition, the smoothness, coherence, and completeness of the essay was important. If the content of the essay merely met the basic requirements of the assigned task and included no further clarifications, it was not given a high score. The expert-assigned scores for the participant essays at the final interpretation stage are shown in the next section, in Table 2. For each task, the five participants with higher report scores were taken as the high-performance (*HP*) group, and the other five with poorer scores were taken as the lower-performance (*LP*) group. We further examined whether learning occurs in the search process as reflected in the final stage of the process (i.e., information synthesis and presentation) based on the learning criteria of Vakkari (2016). The number of learning criteria (*# of Criteria*) met by participants is shown in Table 2. We analyzed the relationship between the report scores and the number of achieved criteria. The learning success criteria adopted in the work and the details of the criteria met by the two groups for the two tasks are shown in the next section, in Table 3.

Code	Observation item	Description
Search formulation		
S	Keyword search	Enter keywords for search
F	Field search	Use the drop-down menu for search
B	Browse search	Switch to browse search and click on the search category
T	Tag cloud search	After clicking the tag, use the tag to search
Selecting sources		
R	Change display settings	Change display settings for search results
C	Click on data	Click on items on search results page
V	View data (SEarch Result Pages, SERPs)	View meta data, size and texture, inscription and imprint, theme technique, reference materials, or pictures
Interacting with sources		
O	Write essays	Paste content, write, and edit essay
Z	Other resources	Search and use resources other than the target DA-PCSS system. , such as: Wikipedia, NPM website, search engines, etc.

Table 1. Code Sheet of Observation Items When Conducting Tasks

PRELIMINARY RESULTS AND DISCUSSION

Discussion on information search process and task performance

Table 2 shows the evaluation results in the different task performance groups for the two tasks. Based on their short essays, in each task we differentiated them into a high-performance (*HP*) group and a lower-performance (*LP*) group, as shown in Table 2. The second column of Table 2 (*Prior K.*) shows the self-assigned user ratings of prior knowledge for the task (the higher the better). The third column (*# of Criteria*) shows the number of learning criteria met by participants based on Table 3 in the next section. Finally, *Scores* indicates the expert-assigned scores for participants' essays at the final interpretation stage.

Task 1—less information support by the system							
HP Group	Prior K.	# of Criteria	Scores	**LP Group**	Prior K.	# of Criteria	Scores
Average	32	9.2	8.6	**Average**	7	5.6	6.4
Task 2 -- more information support by the system							
HP Group	Prior K.	# of Criteria	Scores	**LP Group**	Prior K.	# of Criteria	Scores
Average	35	9.6	8.2	**Average**	17	5.0	6.8

Table 2. Learning Success Criteria Counting Versus Essay Scores

Discussion 1: Better-performing users scored 8.6 on average on the Task 1 report and 8.2 on that for Task 2, whereas lower-performers scored lower than the *HP* group. The latter met fewer search success criteria, such as learning across search stages, than the *HP* group, as shown in Table 2. That is, the *HP* group met 9.4 criteria on average, whereas the *LP* group only met 5.3 criteria. Overall, the evaluation results show that users learn during the search process as reflected both in report scores and in the number of success criteria met derived from search behavior. The notable association between scores for task outcomes and the number of search success criteria met hints that the latter can be used as indicators for the former. Our results suggest that Vakkari's (2016) criteria seem to validly reflect the quality of search outcomes. Furthermore, the prior knowledge in the *HP* groups was notably higher than that in the *LP* group. This reflects the fact that the users' performance in Task 1 did not correlate with the performance in Task 2. The results provide a preliminary response to our research question. Below, we further discuss the characteristics of search and interaction patterns in the three search stages.

Discussion 2 (Formulating searches): In the querying stage, our statistical data shows that the *HP* group entered 11 distinct terms on average whereas the *LP* group entered 7 distinct terms on average for the two tasks. *HP* users submitted more queries and unique keywords than the *LP* users for each task, which suggests that the former were able to formulate keywords to achieve better results. Table 3 shows that nearly all of the users in the *HP* group met the "Increase in number and specificity of terms" and "Increase in number of terms with associative relations (facets) and synonyms" learning criteria. Users in the *HP* group all entered more types of keywords within and outside of the system. In addition to the name of the painting given by the task, they also formulated precise queries that expressed features of Chinese painting, such as "*wrinkle method*" and "*boundary painting*" and painting-related items related such as "*Chinese*", "*Ping tree*", and "*season*". Participant *C*, who achieved better task performance, mentioned that "*I tried to search for information inside the system. But if I wanted to know more about a specific concept, I used a search engine to get its definition, like with the 'wrinkle method' ."* He also remarked that *"I tried not to be limited by the search results when I synthesized the results and wrote the report by myself."* Apparently, he sought out other sources to explore topics in a meaning-making process to link to his prior knowledge of art to write the essay. Participant *I*, with poorer task performance, mentioned, *"I copied phrases provided by the task description, for example, the name of the painting, to search for information using a search engine. Then I returned to the system to search for the painting."* Apparently, she did not know how to select extended keywords and concepts beyond the task description. This shows that most users with better task performance had more prior knowledge of it. In summary, prior domain knowledge seems to play an important role in task performance.

Discussion 3 (Selection sources): In the second stage, *LP* users took more actions—for example, R (Change display settings), C (Click on data), and V (View data) (Table 1) but did not locate the information needed as compared to users in the *HP* group, especially for Task 1, as shown in Figure 3. This also shows that the *LP* group met far fewer learning criteria at this stage than the *HP* group for each task, as shown in Table 3. The *LP* group took more trial-and-error actions because they were exploring unknown subjects with more uncertainty (Borlund & Dreier, 2014). They had less prior knowledge for the two tasks than the *HP* group, as shown in Table 2. Furthermore, Table 3 shows no users in the *LP* group met the "The proportion of specific and factual information increases" learning criterion for both tasks. In contrast, the *HP* group met more learning criteria with fewer actions at this stage, which indicates a better ability to distinguish between relevant information and sources. Figure 3 compares the proportions of interactive actions with associated counts of actions between the two groups for each stage of the tasks.

Discussion 4 (Interacting with sources): In the third stage, we observe that users with better task performance tended to actively seek information outside of the system more frequently. They were more successful during the search formulation stage and interacted with sources more than those in the *LP* group, as shown in Figure 3. The *HP* group took the time to write the essay and verify information iteratively, as reflected in the criteria of Table 3 and the analysis of search paths in the next section. That is, users in the *HP* group all met the "Increase in number and specificity of concepts and their interrelations in the knowledge structure" criterion in Table 3. These results are in accordance with the findings of Borlund and Dreier's (2014) study; that is, the *HP* group used more queries and interacted more with sources (information) since they knew the subject domain better than the *LP* group. As mentioned in earlier studies, this is the most crucial phase for search outcomes. During this phase of result inspection, when users take a longer time or when they take more actions, they do better at completing the task and restructuring task knowledge to produce the outcome (Liu & Belkin, 2012; Vakkari & Huuskonen, 2012). These behaviors suggest that they are making meaning of the assigned task, resulting in higher scores on their reports.

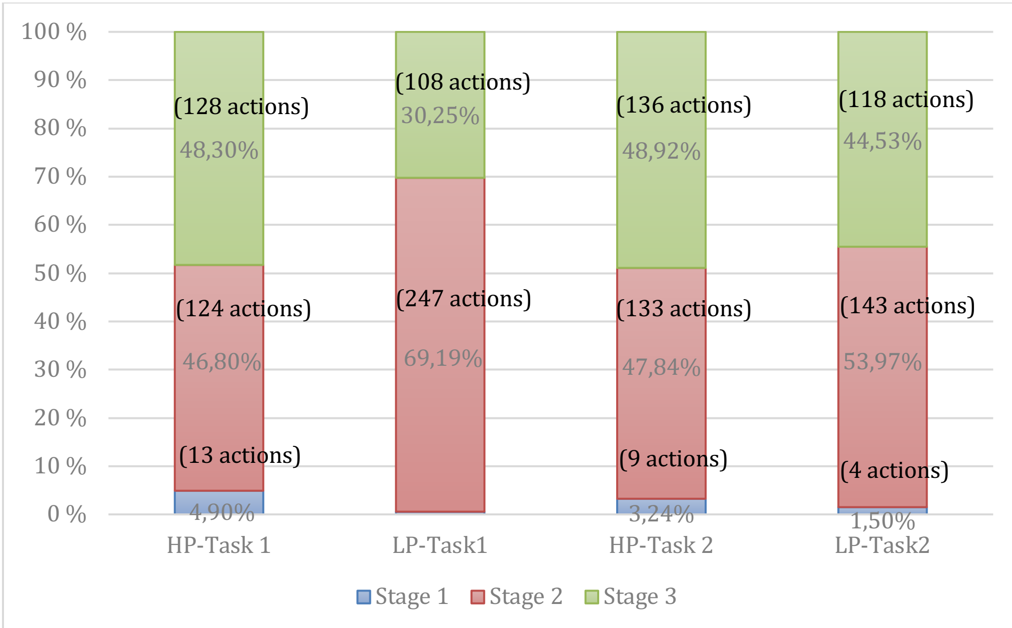

Figure 3. Proportions of Actions for Three Stages of Two User Groups

Search stage	Learning criteria	Task 1		Task 2	
		HP	**LP**	**HP**	**LP**
1ST: Search formulation	Increase in number and specificity of terms	ACGLP	E	KLP	F
	Increase in number of terms with associative relations (facets) and synonyms	ACGLP	EF	ACKL	FG
	Decrease in number of reformulated queries and variability of tactics	AL	E	AL	F
	Decreased time use per search sessions (from keyword search to end of viewing data from SERPs)	ACGP	BFK	CKP	BEI
	Criteria met	**16**	**7**	**12**	**7**
2nd: Sources Selection	Increased clarity in relevance criteria = increased ability to distinguish between relevant and non-relevant sources	ALP	FK	CLP	I
	Decrease in the number of sources viewed in result list	AGLP	FI	K	BEG
	The proportion of sources selected of sources viewed decreases	AC	FK	CLK	E
	The number of sources selected decreases	ACP	EI	LP	B
	The proportion of general background information and theoretical information decreases	AL	KI	ACLKP	F
	The proportion of specific and factual information increases	L		LP	
	Average time used for assessing a source decreases	CP	KI	CLKP	GI
	Criteria met	**17**	**12**	**20**	**9**

3rd: Interaction with sources	Increasing share of sources viewed and selected used in outcome	C	FI	A	EFI
	Increase in number and specificity of concepts and their interrelations in the knowledge structure	ACGLP	EF	ACKLP	FG
	Decrease in the proportion of general background and theoretical information from sources used	C		AKL	G
	Increase in the proportion of specific and factual information utilized from sources for outcome	CGP	FI	AKLP	FG
	Rechecking sources for information initially overlooked	AL	BKI	LP	G
	Criteria met	**12**	**9**	**15**	**9**

Table 3. Learning Criteria Met by Participants within search sessions (Note: Letters denote user IDs)

Discussion on information search process in terms of search patterns

The above evaluation results of search process behaviors were also confirmed by analyzing the results of the ZOST and LSA methods, with which we seek to analyze users search patterns. The number above the line between each rectangle boxes denotes frequency of each search move in Figures 4(a), 4(b), 5(a) and 5(b).

Discussion 1 (ZOSTs): For Task 1, the *LP* group adopted fewer search formulation activities (Codes S, F, B, and T). The *HP* group conducted keyword searches more frequently, as shown in Figures 4(a) and 4(b). The *LP* group used more activities when examining and selecting resources (Codes R, C, and V) compared to the *HP* group. However, compared to the *LP* group, the *HP* group depended more on resources outside of those provided by the database. The ZOST analysis results make it easy to evaluate the impact of system support for different types of tasks; the *HP* group clearly knew where to find knowledge for the task. For Task 1, since the system contains little information related to the "boneless technique", they sought information outside the system (Code Z) to better understand the technique and then write the report (Code O), in comparison to the *LP* group, as shown in Figures 4(a) and 4(b). For Task 2, the *HP* group spent more time within the system because more information was available. Compared to the *LP* group, the *HP* group viewed data more often (Code V) and then worked on the report (Code O), as shown in Figures 5(a) and 5(b). Furthermore, the *HP* group viewed data (V) 2.85 and 1.69 times per addition to the report (O) for Task 1 and 2 respectively, whereas the *LP* group viewed data 1.19 and 1.43 notes per addition for the two tasks. This suggests that the *HP* group explored more information than the *LP* group.

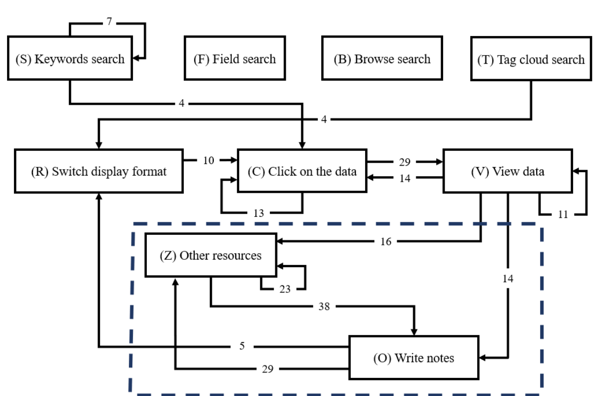

Figure 4(a). ZOSTs for HP group (Task 1)

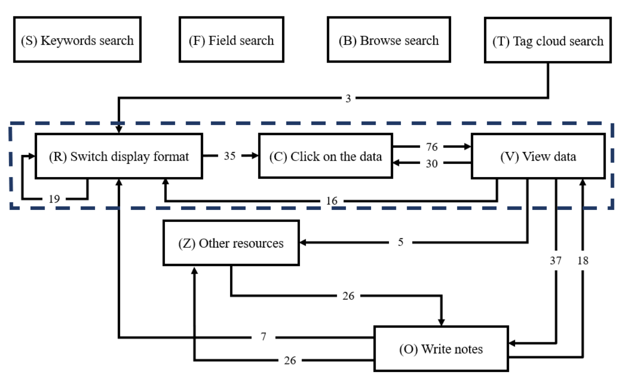

Figure 4(b). ZOSTs for LP group (Task 1)

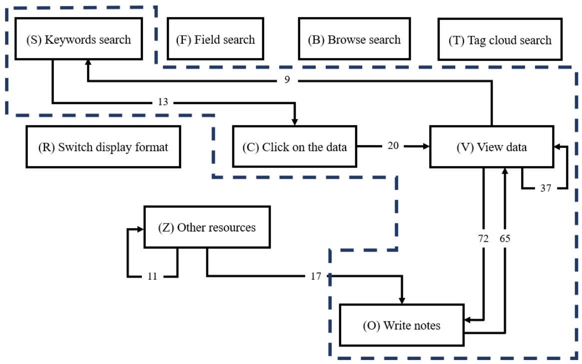

Figure 5(a). ZOSTs for HP group (Task 2)

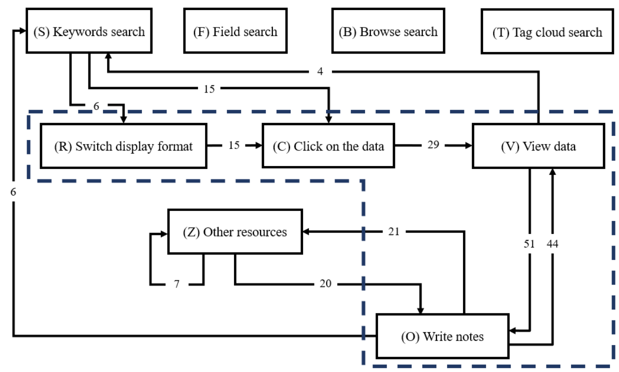

Figure 5(b). ZOSTs for LP group (Task 2)

Discussion 2 (LSA): Here, we focus on the significantly different search path with a length longer than one that are found for the two groups. We list the paths with a 95% significance level, that is, whose z-value is greater than 1.96, based on Bakeman and Gottman's (1997) study, and whose correlation value of the *q* is greater than 0.6. For Task 1, the *HP* group tended to explore information outside the system; that is, they used frequent and significant search paths of O→Z→O for Task 1, as shown in Rule (1), but did not have these for Task 2. For Task 2, the *HP* group had significant search paths of V→O→V and C→V→S, as shown in Rules (2) and (3). This shows that the *HP* group preferred to search and verify information inside the system, and that they were able to find the information needed for better scores. This is in accordance with the analyzed ZOST results. This shows if more information could be retrieved from the system, the *HP* group perceived this and explored and learned more from the system. Figure 6(a) shows the search page of the targeted system. However, the *LP* group did not use these paths. For the *LP* group, it was obvious that they changed the display settings frequently within the system for Task 1: R→C→V, as shown in Rule (4). Figure 6(b), which depicts the webpage for code R of the target system, shows that the *LP* group also spent time on the system to find information. For Task 2, they also used one of the *HP* paths, i.e., O→V→O, as shown in Rule (5). Although they spent more time on the essay, they scored worse than the *HP* group.

Search paths of HP group:

Task 1: O(write essay)->Z(other resources)->O(write essay) {z = 9.159, q= 0.954} (1)
Task 2: V(view SERPs)→O→V {z = 9.107, q = 0.937} (2)
C (click on data)->V->S (keyword search) {z = 2.791, q = 0.620} (3)

Search paths of LP group:

Task 1: R(change display setting)→C→V {z = 9.803, q = 1.000} (4)
Task 2: O→V (view SERPs)→O {z = 6.162, q = 0.781} (5)

Figure 6(a). Keyword Search Page

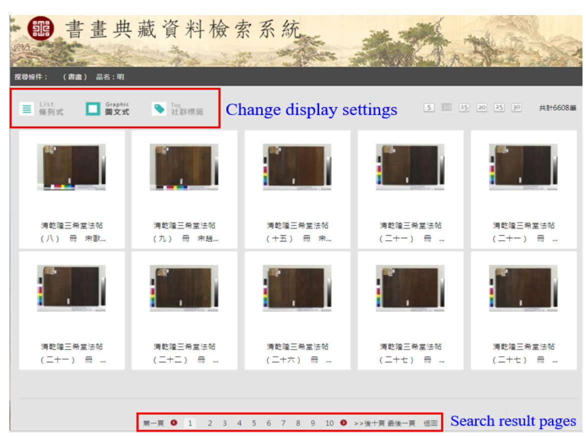

Figure 6(b). SREPs Display Setting Page

CONCLUSION AND FUTURE WORK

In this study we sought to understand how the NPM system fosters users' knowledge of paintings and calligraphy and to examine its effects on their learning. Our findings show that users' with different levels of task performance will have different keywords search and system interaction behaviors, which are all reflected in each search stage. The research results also indicate that learning occurs during searching and that users' task performance is reflected in learning criteria and search behaviors denoting that learning has occurred. Overall, we have better understanding of how the system supports visitors' learning while searching for information online. Furthermore, we found out the participants' prior knowledge seems to explain to a certain extent their search behavior in using the system provided and their task outcome. In line with these results, we will conduct a further investigation into how visitors actively involve themselves in meaning-making by linking their prior knowledge for the task (Skov & Ingwersen, 2014; Skov & Lykke, 2020). Accordingly, we will analyze the results of data collected via the retrospective talk-around research method and the follow-up interview with the aid of a qualitative tool to produce a comprehensive view on the issue. Although there are notable differences between the two groups, the significance of these differences has not been tested. In future work we will include more participants with more types of tasks to validate our empirical results. By examining the effectiveness of the system from the perspective of users' search processes and learning outcomes, this study has furnished a reference for practice.

ACKNOWLEDGEMENTS

This research was supported by the Ministry of Science and Technology, Taiwan under Grant No.108-2410-H-003-132-MY2 and the "Institute for Research Excellence in Learning Sciences" of National Taiwan Normal University (NTNU) from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan. We thank all the volunteers and all publications support.

References

1. Bakeman, R., & Gottman, J.M. (1999). *Observing interaction: An introduction to sequential analysis*. New York: Cambridge University Press.
2. Borlund, P. (2003) The IIR evaluation model: A framework for evaluation of interactive information retrieval systems. *Information Research, 8*(3), http://informationr.net/ir/8-3/paper152.html.
3. Borlund, P. & Dreier, S. (2014). An investigation of the search behavior associated with Ingwersen's three types of information needs. *Information Processing and Management*, *50*(4), 493-507.
4. Chen, P.H., Hsang, j., Chiang, T.M. & Hong, J.S., (2002). Discussion of digital museums, *Museology Quarterly,* 16(3), 15-37.
5. Falk, J. H. (2013). Understanding museum visitors' motivations and learning. In I. B. Lundgaard & J. T. Jensen (Eds.), *Museums: Social learning spaces and knowledge producing processes* (pp. 106-127). Cogenhagen, Denmark: Kulturstyrelsen.
6. Goldman, K. H. & Schaller, D. T. (2004). Exploring motivational factors and visitor satisfaction in on-line museum visits. In Bearman, D. & Trant J. (Eds.), *Museums and the web 2004: Proceedings*. Archives and Museum Informatics: Toronto..
7. Liu J. & Belkin N.J. (2012). Searching vs. writing: Factors affecting information use task performance. In *Proceedings of the American Society for Information Science and Technology*, *49*(1), 1-10.
8. Marty, P. F. (2007) Museum websites and museum visitors: before and after the museum visit. *Museum Management and Curatorship*, *22*(4), 337-360.
9. Pallas, J. & Economides, A. A. (2008). Evaluation of art museums' web sites worldwide. *Information Services & Use*, *28*(1), 45-57
10. Rieh, S.Y., Gwizdka, J., Freund, L. & Collins-Thompson, K. (2014). Searching as Learning: Novel Measures for Information Interaction Research. *Proceedings of the American Society for Information Science and Technology*, *51*(1), 1-4.
11. Ross, C., & Terras, M. (2011). Scholarly information-seeking behaviour in the British museum online collection. In J. Trant & D. Bearman (Eds). *Museums and the Web 2011: Proceedings.* Archives & Museum Informatics: Toronto.
12. Sackett, G.P. (1979). The lag sequential analysis of contingency and cyclicity in behavioral interaction research. In Osofsky, J.D. (Ed.), *Handbook of infant development* (pp. 623–649). New York: Wiley.

13. Sarraf, S. (1999). A survey of museums on the web: Who uses museum websites? *Curator The Museum Journal*, *42*(3), 231-243.

14. Skov, M. (2013). Hobby-related information-seeking behaviour of highly dedicated online museum visitors. *Information Research, 18*(4), paper 597. Available at http://InformationR.net/ir/18-4/paper597.html

15. Skov, M. & Ingwersen, P. (2014). Museum web search behavior of special interest visitors. *Library & Information Science Research*, *36*(2), 91-98.

16. Skov, M., & Lykke, M. (2020). Information-related behaviour as meaning-making processes: a study of science centre visitors. *Proceedings of ISIC, the Information Behaviour Conference, Pretoria, South Africa, 28 September - 1 October, 2020. Information Research, 25*(4).

17. Stebbins, R. A. (2007). *Serious leisure: A perspective for our time*. Transaction Publishers: New Brunswick.

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

19. Vakkari, P. & Huuskonen, S. (2012). Search effort degrades search output but improves task outcome. *Journal of the American Society for Information Science and Technology*, *63*(4), 657-670.

20. Walsh, D., Clough, P., & Foster, J. (2016). User categories for digital cultural heritage. In P. Clough, P. Goodale, M. Agosti, & S. Lawless (Eds.), *First International Workshop on Accessing Cultural Heritage at Scale co-located with Joint Conference on Digital Libraries* (pp. 3-9). Newark, USA.

21. Walsh, D., Hall, M.M., Clough, P., & Foster, J. (2020). Characterising online museum users: a study of the National Museums Liverpool museum website. *International Journal of Digital Library,* 21 (1), 75–87.

22. Wildemuth, B.M. (2004). The effects of domain knowledge on search tactic formulation. *Journal of the American Society for Information Science and Technology*, *55*(3), 246–258