Investigating the Relationship Between Note-Taking Behavior and Learning Outcomes in Learning-Oriented Search: Postprint

Authors: Meng Gaohui, Song Xiaoxuan, Zhang Xiaoyue, Liu Chang

Date: 2023-04-01T15:51:25+00:00

Abstract

[Purpose/Significance] Information search is not merely a process of information acquisition, but can also be viewed as a user's self-learning process. In the context of learning-oriented search, this study investigates the impact of note-taking—an information use behavior during the search process—on learning outcomes, while simultaneously examining the characteristics of users' note-taking behavior and the influence of search context on such behavior. [Method/Process] Data were collected through experimental methods. Encoding analysis, descriptive statistics, and cluster analysis were employed to dissect the characteristics of note-taking, while difference tests were utilized to analyze the associations between note-taking and search context, as well as learning outcomes. [Results/Conclusions] Three note-taking patterns were identified: "coarse collection-linear structure type," "fine collection-linear structure type," and "free structure type." Among these, users with greater prior knowledge tend to adopt the "coarse collection-linear structure type" pattern; however, this pattern may result in poorer learning efficiency. The research conclusions contribute to enriching the limited exploration of information use behaviors in "search as learning" research and inspire the design and optimization of various information search systems from the perspective of supporting users' self-learning.

Full Text

Preamble

Authors: Meng Gaohui[1], Song Xiaoxuan[2], Zhang Xiaoyue[1], Liu Chang[1]
[1]Department of Information Management, Peking University, Beijing 100871
[2]College of Information Management, Nanjing Agricultural University, Nanjing 210095

Abstract: [Purpose/Significance] Information search is not only a process of information acquisition but also a process of users' self-learning. In the context of learning-related search, this study explores the influence of note-taking—an information use behavior during the search process—on learning outcomes, while also examining the characteristics of users' note-taking behavior and the impact of search context on note-taking behavior. [Method/Process] Data were collected using experimental methods, with note-taking characteristics analyzed through coding analysis, descriptive statistics, and cluster analysis. Differences in note-taking behavior across search contexts and learning outcomes were examined using difference tests. [Result/Conclusion] Three note-taking patterns were identified: "Rough Collection-Linear Structure," "Fine Collection-Linear Structure," and "Free Structure." Users with more prior knowledge tend to adopt the "Rough Collection-Linear Structure" pattern, but this pattern may lead to poorer learning efficiency. The findings help enrich the limited exploration of information use behavior in "search as learning" research and inspire the design and optimization of information search systems from the perspective of supporting users' self-learning.

Keywords: Note-taking; Learning-related search; Search as learning; Learning outcome; Knowledge level
Classification Number: G251
DOI: 10.13266/j.issn.0252-3116.2022.08.005

Information search systems have become indispensable tools for people seeking information. Traditional search systems have focused primarily on information retrieval and selection. However, as society has entered the knowledge era, information search is not merely a simple information acquisition process but can also be viewed as a user's self-learning process. Search systems can be further designed and optimized from the perspective of supporting users' self-learning, transforming them into learning systems that support information searching, browsing, understanding, analyzing, and creating [1]. Against this backdrop, the theme of "searching as learning (SAL)" has received continuous attention. Note-taking is one of the most common activities during search sessions. Users' note-taking during search represents a complex cognitive process of transforming perception into association, analysis, and synthesis, and then into written expression [4], which can be regarded as an information use behavior involving the comprehension and processing of acquired information. Based on this, in the context of learning-related search, this study investigates the influence of note-taking—an information use behavior—on learning outcomes during the search process, aiming to enrich the limited exploration of information use behavior in SAL-related research.

Literature Review

2.1 Research on Factors Influencing Learning Outcomes in the "Search as Learning" Context

Scholars in the information search field believe that future search system design should shift from answer-centric to learning-centric, creating good knowledge contexts for searchers to serve their long-term memory [5]. SAL-related research combines theories from learning science to study information search, viewing search systems as learning tools [6], focusing on the role of search systems in learning processes and outcomes, and optimizing, expanding, and reconfiguring search system features and functions to transform them from information retrieval tools into learning spaces that serve human learning [7].

In the SAL context, research on information behavior has primarily focused on how query formulation and reading-related behaviors affect learning outcomes. However, the search process with learning as its goal is complex; users must not only search for and read information but also understand, select, compare, write, and create information [2]. Existing research has limited exploration of the relationship between subsequent information use behavior and learning outcomes [3]. Note-taking behavior in search can be seen as an information use behavior that integrates newly acquired information with users' existing knowledge structures, organizing, expressing, and outputting learned knowledge. Currently, SAL-related research on information behavior has concentrated on how query formulation and reading behaviors affect learning outcomes, with limited exploration of the relationship between subsequent information use behavior and learning outcomes [3].

2.2 Impact of Note-Taking on Learning Outcomes in Classroom Settings

Note-taking means that users comprehend a written document or speech and record information by writing it down [13]. Note-taking is a manifestation of cognitive ability in learning activities under the influence of metacognition and resource management strategies [14] and forms the primary basis for educational activities between teachers and students [15]. Over the past decades, educational research has extensively studied the learning effects of note-taking in classroom settings, forming two hypotheses about note-taking functions: the storage function hypothesis and the encoding function hypothesis [16-17]. The storage function hypothesis posits that the effect of note-taking arises from possession of the notes; reviewing notes can trigger re-cognition of lecture content and consolidate learning. The encoding function hypothesis argues that the effect of note-taking arises from the information processing itself; the note-taking process can enhance learners' attention, trigger active learning activities, and lead to more elaborate thinking, which benefits memory organization and transfer. The storage function of note-taking has been confirmed, but findings on the encoding function are mixed, as different learners' note-taking strategies and

quality may affect the degree to which the encoding function is realized [18].

Existing research has classified note-taking strategies along different dimensions and measured note quality to explore its impact on learning effects. For example, K.A. Kiewra et al. divided note-taking techniques into traditional, linear, and matrix techniques. Traditional techniques produce free notes without specific structure; linear techniques produce outline notes with hierarchical and list-based structures; matrix techniques produce matrix notes in the form of two-dimensional cross-tables that allow cross-category comparisons horizontally and vertically. This research found that under both immediate test conditions (where participants could not review notes) and delayed test conditions (where participants could review notes), matrix note-takers performed better on recall tests than free note-takers [19]. Further, K.A. Kiewra and S.L. Benton found that linear notes were more beneficial than other formats for improving essay writing test scores [20]. Hu Jin also compared these three recording strategies, finding that under appropriate knowledge control conditions, the generation processes for linear and matrix techniques occurred during note-taking, while the generation process for traditional techniques could only occur later during note review, with the former showing significantly better generative effects than the latter [21].

M. Pottbast and M. Hagen et al. identified two different text recording strategies based on the length of text completed after each editing session: buildup and boil-down. Buildup refers to writers continuously increasing text length throughout the recording process; boil-down refers to writers first accumulating substantial material with rapidly increasing text length in the early stage, then reorganizing existing text to shorten it [22-23]. J. Wiley and J.F. Voss analyzed the quality of historical essays written by participants across dimensions including narrative organization, completeness, information sources, scope of explanation, and connections between conceptual units [24]. Regarding information sources, sentences in essays were coded as "borrowed," "transformed," or "added" [24-26]. Similarly, M. Hagen et al. coded note-taking units based on content transformation and integration degree as "paraphrases," "intratextual elaborations," or "intertextual elaborations," where paraphrasing involves little or no content transformation and represents a shallow processing strategy, while elaboration involves more content transformation and represents a deeper processing strategy [27]. They also measured the number of texts covered, number of switches, and number of words in notes, referring to how many documents were mentioned, how many times notes switched between documents, and the word count in notes, respectively. Analyzing the relationship between these note-taking variables and text comprehension performance, they found that more intertextual elaborations in notes correlated with better deep comprehension of single and multiple texts [27].

Overall, these studies classified and measured note-taking strategies and quality along dimensions such as linear recording technique, length accumulation patterns, length, generation degree, organization degree, coverage degree, and

integration degree, finding that different strategies and qualities produce different learning effects.

2.3 Impact of Note-Taking on Learning Outcomes in Search Scenarios

In addition to classroom settings, note-taking is also common in web search scenarios. Note-taking is one of the most frequent activities in search sessions, with users employing handwritten and electronic notes to collect, manage, organize, and share information found during online searches [28]. While numerous studies have examined the impact of note-taking on learning outcomes in classroom settings, research on its impact in web search scenarios has been limited until recently.

X. Song et al. evaluated users' learning outcomes using seven indicators extracted from post-search task essays, finding that note-taking behavior had a significant effect on learning outcomes, with earlier information recording leading to better learning outcomes [29]. J.Y. Wu and C. Xie evaluated learning outcomes using a series of true/false and subjective questions after searching, finding that users under the matrix note strategy performed better than those under no-note and free-note strategies [30]. Later, J.Y. Wu also found that free notes and matrix notes could effectively help users improve learning outcomes under different individual working memory capacities [31]. Liu Chang et al. evaluated users' recording effectiveness and learning outcomes by the number of topical facets covered in notes and classified users' note-taking strategies into early-recording, average-recording, and late-recording types based on recording completion progress over total task time, finding that the average-recording type produced the most facets [32]. N. Roy et al. evaluated users' learning outcomes in receptive learning tasks using vocabulary knowledge tests and in critical learning tasks using essay writing quality, finding that users with note-taking tools included more factual knowledge in their essays [33]. However, these studies only measured or classified note-taking behavior from limited dimensions, focusing on the presence or absence of notes or the form and format of note-taking, without examining how note content quality affects learning outcomes.

In summary, the search process with learning as its goal is complex. Searching and reading are only partial behaviors; subsequent information use behaviors—understanding acquired information, integrating it with personal knowledge structures, and organizing, expressing, and outputting learned knowledge—are also important. Currently, SAL-related research on information behavior has focused on how query formulation and reading behaviors affect learning outcomes, with limited exploration of the relationship between subsequent information use behavior and learning outcomes [3]. Users' note-taking behavior in search can be viewed as an information use behavior that integrates new information with existing knowledge structures, representing a key step in the learning process. Compared to classroom settings, research on note-taking and its impact on learning outcomes in search scenarios is limited. A few relevant studies have only measured or classified note-taking behavior from partial dimensions, focusing on

the presence or form/format of notes without examining how note content quality affects learning outcomes. Based on this, this study measures and classifies note-taking behavior from multiple dimensions including both form and content, deeply exploring the impact of note-taking—an information use behavior—on learning outcomes in learning-related search contexts, while also examining the characteristics of users' note-taking behavior and the influence of search context on note-taking behavior, thereby enriching the limited exploration of information use behavior in SAL research and inspiring search systems to design and optimize functions from the perspective of supporting users' self-learning.

Data Collection Methods and Results

3.1 User Experiment

3.1.1 Participants This study utilized experimental data from a previous project; specific experimental design and procedures reference Song Xiaoxuan and Liu Chang (2018) [34]. The experiment recruited 32 participants from Peking University, including 12 males and 20 females aged 18-30. Participants came from different departments: 14 from humanities and social sciences, 5 from information and engineering sciences, 10 from natural sciences, and 3 from medical sciences. They were at different academic levels: 20 undergraduates (8 freshmen, 3 sophomores, 4 juniors, 5 seniors), 4 master's students, and 8 doctoral students. The experiment was conducted from March to April 2016, with an average duration of 1.5 hours.

3.1.2 Experimental Procedure Before the experiment, participants completed the Extended Cognitive Style Analysis: Wholist-Analytic test (E-CSA-WA) [35] and a background questionnaire. During the experiment, participants performed searches on four search tasks sequentially. Before each task, participants completed a pre-search questionnaire rating their task familiarity and perceived task difficulty, and described in writing what they already knew about the task. They then conducted the search without time limits, being asked to record and organize what they considered reasonable answers in a computer notepad file. They could stop searching at any time when they felt they had completely recorded their answers. After searching, participants completed a post-search questionnaire rating their post-search task familiarity and perceived task difficulty, and described their answers to the task in writing without referring to external information. After completing all four tasks, participants filled out a post-experiment questionnaire providing suggestions for search system design. Throughout the experiment, Morae Recorder 3.3 was used to record users' search interaction behaviors in the background. The experimental procedure is shown in Figure 1 [Figure 1: see original paper].

3.1.3 Search Task Design The search tasks used simulated learning-related search tasks. Learning-related search tasks refer to search activities where users, in a learning context and with learning as the goal, use search systems to col-

lect, analyze, evaluate, and use information to ultimately complete the learning task [32, 34, 36]. Analyzing user interaction behaviors in learning-related search tasks can maximize the externalization of relatively implicit learning processes. Therefore, this study adopted learning-related search tasks in the experiment, using S.Y. Rieh et al.'s [37] learning process classification as the theoretical basis for task design. Tasks focused on daily health issues of concern to college students, designing two types: receptive learning tasks and critical learning tasks, with two different search topics for each type, totaling four tasks. Receptive learning refers to students' memory, understanding, and expression of learned knowledge—a relatively shallow learning process. Critical learning refers to students' reflection, criticism, and evaluation of knowledge from different sources to form their own thinking—a deeper learning process than receptive learning. Each participant completed all four tasks, with task order rotated across participants using a Latin square method.

The four search tasks were:

(1) Receptive Learning Tasks:

Task 1: Beijing's smog is becoming increasingly severe, and you feel it is necessary to wear a mask during smoggy weather. Before buying a mask, you want to know what types of masks exist, the differences between them, and which type is suitable for young people to wear long-term in Beijing. Please search online and record your answers in the notepad during the search process.

Task 2: An elderly family member has hypertension and needs to purchase a blood pressure monitor for regular monitoring. However, the family lacks experience in buying blood pressure monitors. Please search online and describe what aspects should be considered when evaluating the quality of home blood pressure monitors, recording your answers in the notepad during the search process.

(2) Critical Learning Tasks:

Task 3: Online, there are claims that snow boots have adverse health effects, causing problems with feet, ankles, and hips. At the same time, many netizens have refuted these claims. Please analyze and compare the two viewpoints and summarize your own opinion on whether snow boots affect health. Search online and record your answers in the notepad during the search process.

Task 4: Your middle school cousin is considering joining the school soccer team. Most relatives support this idea, but some argue that soccer is a dangerous sport with potential health risks. His parents also want your opinion: What are the health hazards of long-term soccer playing for adolescents? Should they let the child join the soccer team? Please search online and record your answers in the notepad during the search process.

3.2 Search Context and Learning Outcome Variables

The search context and learning outcome variables measured in this study are shown in Table 1 . Some variables are explained below:

- **Cognitive Style:** Measures people's habitual ways of thinking, viewing, and constructing information. Wholist style refers to people who tend to view the entire situation as a whole and try to obtain as complete information as possible. Analytic style refers to people who tend to break down situations into parts and focus on one or several aspects [38-39].
- **Knowledge Level:** This study comprehensively evaluated participants' knowledge levels from quantitative and qualitative dimensions. Quantitative indicators included word count, knowledge point count, knowledge facet count, knowledge breadth, and knowledge depth. Qualitative indicators included objective knowledge relevance, knowledge analysis degree, subjective task familiarity, and perceived task difficulty. Based on participants' written descriptions of task knowledge before and after searching, this study used coding analysis to obtain values for knowledge level variables before and after searching. The process is described in existing literature [34], with pre-search knowledge level treated as a personal characteristic and post-search knowledge level treated as a learning outcome.

3.3 Note-Taking Variables

3.3.1 Variable Selection The note-taking variables measured in this study are shown in Table 2 , including two categories: note-taking strategies and note quality.

(1) Note-Taking Strategies: Judged from a dynamic perspective regarding the strategies participants used during note-taking, including: - **Linear Technique:** Measures participants' tendency to use list-based, hierarchical linear recording techniques - **Accumulation Pattern:** Measures the length accumulation pattern of participants' note content, including buildup and boil-down types [22-23], coded as 1 and 2 respectively

(2) Note Quality: Judged from a static perspective regarding the quality of the final notes produced, including: - **Length:** Measures note length, including word count, unit count, and average unit length - **Generation Degree:** Different note-taking procedures have different generation degrees [18]. Note-taking procedures can be generative or non-generative [40]. Generative procedures represent deep processing of original information involving relatively deep cognitive processes, while non-generative procedures represent shallow processing involving superficial cognitive processes [41-42]. The "generation degree" variable measures the generation degree of note-taking procedures. - **Processing Degree:** Measures the depth of note processing of original web information, including text selection degree, content transformation degree, and position reorganization degree - **Organization Degree:** Measures the degree of note organization

and structuring - **Coverage Degree and Integration Degree:** Measures the degree to which participants' notes covered and integrated different information sources

3.3.2 Coding Analysis Five note-taking variables were obtained through coding analysis: "Linear Technique," "Text Selection Degree," "Content Transformation Degree," "Position Reorganization Degree," and "Organization Degree." Coding rules drew on existing literature methods and conclusions and were formulated based on this study's data.

(1) "Linear Technique" Variable Coding: This involved two-level coding. In the first level, each action unit where participants added linear formats during note-taking served as the coding unit, with three features coded: operation level (which text level the action targeted, coded as N where N ranged 1-3 in this study), specific operation (type of linear format added, coded as adding subtitles or adding bullet points), and operation scope (completeness of linear format added to the text at that level, coded as partial, some, or whole). Based on first-level coding results, the second level used action unit collections within the same operation level as coding units. The linear technique used at each level was coded as no outline, basically no outline, basically has outline, or has outline, assigned values of 0, 1, 2, and 3 respectively, with higher scores indicating more complete linear recording techniques at that level. Values for all levels were summed to obtain the note's score.

(2) "Text Selection Degree," "Content Transformation Degree," and "Position Reorganization Degree" Variables Coding: This study defined units as spatially proximate, content-related, and informationally complete text segments in notes [27], generally consisting of one sentence. Coding analysis used each unit as the coding unit, evaluating three dimensions: - **Text Selection Degree:** The fineness of original web text selection, coded as no selection, paragraph selection, sentence selection, or word selection, assigned values of 0, 1, 2, and 3 respectively, with higher scores indicating finer selection - **Content Transformation Degree:** The degree of processing of original web text content, with coding shown in Table 3 , where higher scores indicate deeper processing - **Position Reorganization Degree:** The degree of reorganizing original web text positions, coded as no change, position swapping, or special positioning, assigned values of 0, 1, 2, and 3 respectively, with higher scores indicating more reorganization

Each unit was coded and assigned values, then averaged across all units in a note to obtain the note's scores.

(3) "Organization Degree" Variable Coding: This coding analysis used each text level in notes as the coding unit, coding based on relationship categories, organization forms, and organization scope at that level. Relationship categories refer to relationships between text parts at that level, organization forms refer to types of organizational formats added, and organization scope

refers to completeness of organizational formats added. Each level's organization degree was coded as no organization, basically no organization, basically has organization, or has organization, assigned values of 0, 1, 2, and 3 respectively, with higher scores indicating higher organization and structuring at that level. Values for all levels were summed to obtain the note's score.

As this was an exploratory study, coding initially selected two search tasks (Tasks 1 and 3). One coder performed manual coding of note files, then a second coder independently coded a random sample of 8 notes (13%, meeting the minimum 10% sample size requirement for inter-coder reliability testing [43]). Inter-coder reliability was assessed: Cohen's Kappa for "Text Selection Degree" = 0.760 ($p < 0.001$), for "Content Transformation Degree" = 0.718 ($p < 0.001$), and for "Position Reorganization Degree" = 0.706 ($p < 0.001$), all above 0.7 indicating good agreement. Additionally, Kappa for "Linear Technique" = 0.880 ($p < 0.001$) and for "Organization Degree" = 0.826 ($p < 0.001$), both above 0.8 indicating high agreement. During coding, one search session's notes were excluded due to missing interaction behavior data. Data on note-taking variables were ultimately collected for 63 notes from 32 participants, forming the basis for subsequent analysis.

Data Analysis Process and Results

Data analysis proceeded in three stages: first analyzing characteristics of user note-taking behavior, then examining associations between search context and note-taking behavior, and finally analyzing associations between note-taking behavior and learning outcomes. SPSS software (version R24.0.0.0) was used for all analyses.

4.1 Characteristics of User Note-Taking Behavior

4.1.1 Descriptive Statistics of Note-Taking Behavior Descriptive statistics for all note-taking variables are shown in Table 4 .

Regarding note-taking strategies, linear technique showed low right skew and flat kurtosis, indicating most participants slightly favored traditional recording techniques, with varying degrees of linear technique use across notes. Among all notes, 54 (86%) adopted the buildup accumulation pattern, while only 9 (14%) adopted the boil-down pattern, indicating that participants more commonly accumulated text content progressively, with only a minority adopting a "collect first, refine later" approach.

Regarding note quality, word count and unit count showed high right skew and high kurtosis, indicating most notes had relatively few words and units, with a few notes having large volumes and very concentrated distributions. Average unit length showed a normal distribution, indicating different notes' average unit lengths followed a normal pattern. Text selection degree showed low left skew and flat distribution, indicating slightly more notes made relatively fine

selections of original web text, with varying selection fineness across notes. Content transformation degree showed moderate right skew and flat distribution, indicating more notes did not deeply process original web content, with varying processing degrees across notes. Position reorganization degree showed high right skew and peaked distribution, indicating most notes did not reorganize web text positions at scale, with most units maintaining original positions and concentrated distributions. Organization degree showed low right skew and peaked distribution, indicating more notes had low organization and structuring, with free structures and concentrated distributions. Coverage degree showed high right skew and high kurtosis, indicating most notes involved few information sources with very concentrated distributions. Integration degree showed high right skew and peaked distribution, indicating most notes did not deeply integrate information sources, with concentrated distributions.

4.1.2 User Note-Taking Patterns Cluster analysis was performed based on 10 continuous variables from note-taking variables to explore user note-taking patterns. Before clustering, exploratory factor analysis was conducted. The KMO value was 0.560 (>0.5), and Bartlett's test of sphericity was significant ($p < 0.001$), indicating factor analysis was appropriate. The sample-to-variable ratio was 6.3, exceeding the reference standard of 4 [44], suggesting good factor analysis results.

Principal component analysis with varimax rotation extracted three factors with eigenvalues >1, explaining 80.369% of variance, exceeding the 40% standard [45]. After rotation, item loadings ranged from 0.716 to 0.934, all exceeding the minimum requirement of 0.5, as shown in Table 5 .

Factor 1: Higher values represent longer notes with longer average unit lengths, and coarser selection and processing of original web text (direct large-scale collection). This factor was named "Rough Collection Degree."

Factor 2: Higher values represent notes covering more information sources with deeper integration and more extensive reorganization of original web text positions. This factor was named "Multi-Source Integration Degree."

Factor 3: Higher values represent higher organization degrees in notes and greater use of linear recording techniques. This factor was named "Linear Organization Degree."

Factor scores from SPSS were saved. For easier interpretation, Factor 1 was reversed and renamed "Fine Collection Degree," where higher values indicate finer selection and processing of original web text and shorter lengths and unit lengths.

Based on these three factors, K-means clustering was applied to the 63 notes. After adjustment, K=3 was selected based on "Fine Collection Degree" and "Linear Organization Degree," dividing notes into three clusters. One-way ANOVA showed significant differences across the three clusters on both variables ($p <$

0.001), indicating valid clustering results. Table 6 describes the three patterns:

(1) Rough Collection-Linear Structure (RC-LS): n=12 (19%). This pattern had far lower fine collection degree (-1.276) than the other two patterns but higher linear organization degree (0.919). Notes made coarse selections of original web text without deep processing, using shallow processing strategies, but had longer lengths and unit lengths. Text was highly organized with extensive use of list-based linear recording techniques.

(2) Fine Collection-Linear Structure (FC-LS): n=15 (24%). This pattern had far higher fine collection degree (0.815) than the other two patterns and the highest linear organization degree (1.006). Notes made fine selections of original web text with deep processing, had shorter lengths and unit lengths, and featured highly organized text with extensive use of list-based linear recording techniques.

(3) Free Structure (FS): n=36 (57%), the most frequent pattern. This pattern had moderate fine collection degree (0.086) but far lower linear organization degree (-0.725) than the other two patterns. Notes made moderate selections and processing of original web text, had medium length and short unit lengths, but featured low organization with free structures. Participants primarily used traditional recording techniques, simply recording notes in the order text was encountered during tasks.

4.2 Association Between Search Context and Note-Taking Behavior

Associations between search context variables and note-taking patterns were examined, with significant results shown in Tables 7 and 8 .

First, the three patterns differed significantly on multiple variables reflecting pre-search knowledge level, including pre-search word count ($p = 0.030$), pre-search knowledge point count ($p = 0.006$), pre-search knowledge facet count ($p = 0.001$), pre-search knowledge breadth ($p = 0.004$), pre-search knowledge relevance ($p = 0.010$), pre-search task familiarity ($p = 0.039$), and pre-search perceived task difficulty ($p = 0.028$). Pairwise comparisons showed that RC-LS > FS was significant for word count, knowledge point count, knowledge facet count, knowledge breadth, knowledge relevance, and task familiarity, and RC-LS > FC-LS was also significant for knowledge relevance. This indicates that participants adopting the RC-LS pattern, compared to those adopting the other two patterns (especially FS), had more prior knowledge, including more knowledge points, broader knowledge facets, and greater task familiarity.

Second, note-taking patterns were significantly associated with task type (Pearson chi-square test, $p = 0.014$). Comparing actual versus expected counts, in receptive learning tasks, FC-LS actual counts were significantly higher than expected, while FS actual counts were significantly lower than expected. In critical learning tasks, FC-LS actual counts were significantly lower than expected, while FS actual counts were significantly higher than expected. This indicates

participants were more likely to adopt the FC-LS pattern in receptive tasks and less likely to adopt the FS pattern, with the opposite pattern in critical tasks.

4.3 Association Between Note-Taking Behavior and Learning Outcomes

Associations between note-taking patterns and learning outcome variables were examined, with significant results shown in Table 9 . The three patterns did not differ significantly on variables reflecting post-search knowledge level, but showed significant differences on two variables reflecting knowledge learning efficiency: content mastery rate ($p = 0.004$) and knowledge facet mastery rate ($p = 0.002$). Pairwise comparisons showed that FC-LS > RC-LS and FS > RC-LS were both significant for both content mastery rate and knowledge facet mastery rate. This indicates that participants adopting the RC-LS pattern had lower learning efficiency from notes, including lower mastery of note content and fewer knowledge facets mastered per unit, while the other two patterns did not differ significantly in learning efficiency.

Discussion

Under learning-related search conditions, this study addressed three questions: characteristics of users' note-taking behavior during search, associations between search context and note-taking, and associations between note-taking and learning outcomes.

First, this study measured note-taking behavior from multiple dimensions including both form and content, revealing overall characteristics of users' recording behaviors. Three note-taking behavior patterns were identified: "Rough Collection-Linear Structure," "Fine Collection-Linear Structure," and "Free Structure." The first two patterns share highly organized text with extensive use of list-based linear recording techniques, differing in that RC-LS involves direct large-scale collection of original web text, while FC-LS involves fine selection and deep processing. The FS pattern differs primarily in its low organization and free structure, with participants primarily using traditional recording techniques in the order text was encountered. This pattern was most frequently used, likely because traditional techniques are simple, align with user habits, and allow following the search flow [46].

Second, this study examined associations between search context and the three note-taking patterns. Regarding personal characteristics, users with more prior knowledge tended toward the RC-LS pattern, while those with less prior knowledge favored the FS pattern. This aligns with existing literature showing that incomplete topic understanding leads to structural uncertainty and that lack of topic knowledge affects complex structure use [46]. No association was found between cognitive style and note-taking behavior, consistent with findings that cognitive style primarily affects search behavior while topic familiarity mainly affects recording behavior [39]. Regarding task characteristics, participants were

more likely to adopt the FC-LS pattern in receptive learning tasks and less likely to adopt the FS pattern, with the opposite in critical learning tasks. This may relate to prior knowledge effects, as participants in this study had significantly higher pre-search knowledge levels for receptive tasks.

Third, this study examined associations between the three note-taking patterns and learning outcomes. The finding that the RC-LS pattern led to lower learning efficiency aligns with generative theory. Notes with high generation degrees promote learners' active integration of new knowledge with existing knowledge, paraphrasing, organizing, and finding meaning in new material [47], thereby deepening thinking, strengthening summarization, optimizing perception, maintaining attention, and facilitating memory [14]. Under learning-related search conditions, load theory of attention and cognitive control further explains this generative theory. On one hand, search tasks occur in rich internet environments full of irrelevant distracting information, posing great challenges to attention concentration. As an active reading strategy, note-taking can strengthen learners' perception in selecting and extracting relevant information, increase active attentional selection, and inhibit task-irrelevant stimuli [30-31]. On the other hand, learning-related search tasks involve complex information processing and decision-making, requiring extensive higher-order cognitive abilities. When reading and comprehending new information, note-taking helps learners externalize thoughts and express them clearly, thereby controlling cognitive load [33]. Therefore, the RC-LS pattern with low generation degree led to lower learning efficiency compared to the other two patterns with higher generation degrees.

These findings reveal an interesting phenomenon: users with more prior knowledge tended to adopt the RC-LS pattern, which ultimately led to lower learning efficiency. Users with the highest prior knowledge, focusing on large-scale information collection and linear text organization while neglecting selection and processing of original web information, achieved the worst learning efficiency. This reflects the Dunning-Kruger effect, where users often overestimate their knowledge levels after obtaining information through search engines, mistakenly believing that knowledge acquired from the internet exists in their own minds [48]. In reality, they have not truly learned this knowledge; the ease of information access creates this illusion.

Conclusion

This study explored the impact of note-taking behavior on learning outcomes in learning-related search processes, while analyzing characteristics of note-taking behavior and the influence of search context. First, while SAL-related research has focused on how query formulation and reading behaviors affect learning outcomes, this study examined note-taking as an information use behavior, enriching limited exploration in this area. Second, previous note-taking research has primarily been in educational classroom settings, with limited attention to search scenarios. This study examined note-taking in search contexts from multiple dimensions including form and content, 弥补了已有研究中关注维度有限

的缺陷，将教育学中的研究方法应用于搜索行为研究，对搜索行为研究方法进行拓展和创新。
Finally, note-taking behavior can explicitly reflect when and how learning processes occur during search, externalizing relatively implicit learning processes. This study's findings further extend and deepen current understanding of user learning processes.

The study also has practical value. The findings identify user note-taking patterns and their associations with search context and learning outcomes, inspiring search systems to improve note-taking services from a learning support perspective. For example, systems could monitor users' note-taking patterns to infer search context and learning efficiency, or recommend appropriate note-taking patterns based on search context to guide users toward patterns that enhance learning efficiency, thereby supporting effective learning during search.

Due to time and resource constraints, this study has limitations. First, participants were limited to Peking University students, which may affect generalizability. Future research should expand the sample range to test result robustness. Second, some variables were obtained through coding analysis, which involves subjectivity, and only 13% of samples were used for reliability testing, which may affect variable reliability. Future research should sample more extensively to further verify coding reliability. Finally, as an exploratory study, this paper only provides preliminary insights for improving note-taking services in search systems; future research should explore specific methods and tools for implementing such improvements in actual search systems.

References

[1] Meyer A, Hansen P, Fourie I. Assessing the potential of third space to design a creative virtual academic space based on findings from information behavior. Information research, 2018, 23(2): paper isic1814. http://www.informationr.net/ir/23-4/isic2018/isic1814.html.

[2] Vakkari P. Task-based information searching. Annual review of information science & technology, 2003, 37(1): 413-464.

[3] Song Xiaoxuan, Liu Chang, Chen Jianlong. Review of research on search as learning. Library and Information Service, 2021, 65(10): 113-126.

[4] Hu Jin. On innovative classroom note-taking strategies. Shanghai Educational Research, 2000(3): 58-59.

[5] Smith CL, Rieh SY. Knowledge-context in search systems: toward information-literate actions. Proceedings of the 2019 conference on human information interaction and retrieval. New York: ACM, 2019: 55-62.

[6] Rieh SY, Gwizdka J, Freund L, et al. Searching as learning: novel measures for information interaction research. Proceedings of the Association for Information Science and Technology, 2014, 51(1): 1-4.

[7] Hansen P, Rieh SY. Recent advances on searching as learning: introduction to the special issue. Journal of information science, 2016, 42(1): 3-6.

[8] Tan Jinbo. Impact of students' implicit and explicit information search strategies on web-based learning performance. China Educational Technology, 2014(9): 67-71, 77.

[9] Gwizdka J, Chen X. Towards observable indicators of learning on search. Proceedings of SIGIR2016 workshop: searching as learning (SAL), 2016[2022-01-30]. http://ceur-ws.org/Vol-1647/SAL2016_{{{paper}}}_{{19}}}.pdf.

[10] Bhattacharya N, Gwizdka J. Relating eye-tracking measures with changes in knowledge on search tasks. Proceedings of the 2018 ACM symposium on eye tracking research & applications. New York: ACM, 2018: 1-5.

[11] Collins-Thompson K, Rieh SY, Haynes CC, et al. Assessing learning outcomes in web search: a comparison of tasks and query strategies. Proceedings of the 2016 conference on human information interaction and retrieval. New York: ACM, 2016: 163-172.

[12] Gadiraju U, Yu R, Dietze S, et al. Analyzing knowledge gain of users in informational search sessions on the web. Proceedings of the 2018 conference on human information interaction & retrieval. New York: ACM, 2018: 2-11.

[13] Piolata A, Oliver T, Kellogg RT. Cognitive effort during note-taking. Applied cognitive psychology, 2005, 19(3): 291-312.

[14] Fan Sufang, Fan Qi, Chen Jie. Current status of college students' classroom note-taking strategies. Psychological and Behavioral Research, 2007, 5(1): 70-74.

[15] Castellà M, Monereo C. Students' note-taking as a knowledge-construction tool. Educational studies in language and literature, 2005, 5(3): 265-285.

[16] Hagverdi H, Biria R, Karimi L. Note-taking strategies and academic achievement. Journal of language and linguistic studies, 2010, 6(1): 75-109.

[17] Kiewra KA. A review of note-taking: the encoding-storage paradigm and beyond. Educational psychology review, 1989, 1(2): 147-172.

[18] Kobayashi K. What limits the encoding effect of note-taking? A meta-analytic examination. Contemporary educational psychology, 2005, 30(2): 242-262.

[19] Kiewra KA, DuBois NF, Christian D, et al. Note-taking functions and techniques. Journal of educational psychology, 1991, 83(2): 240-245.

[20] Kiewra KA, Benton SL. Effects of note-taking format and study technique on recall and relational performance. Contemporary educational psychology, 1995, 20(2): 172-187.

[21] Hu Jin. Experimental study on college students' note-taking strategies. Psychological Science, 1999, 22(4): 377-378.

[22] Pottbast M, Hagen M, Völske M, et al. Crowdsourcing interaction logs to understand text reuse from the web. Proceedings of the 51st annual meeting of the association for computational linguistics. Pennsylvania: The Association for Computational Linguistics, 2013: 1212-1221.

[23] Hagen M, Pottbast M, Gomoll J, et al. How writers search: analyzing the search and writing logs of non-fictional essays. Proceedings of the 2016 ACM conference on human information interaction and retrieval. New York: ACM, 2016: 193-202.

[24] Wiley J, Voss J. The effects of "playing historian" on learning in history. Applied cognitive psychology, 1996, 10(7): 63-72.

[25] Wiley J, Voss JF. Constructing arguments from multiple sources: tasks that promote understanding and not just memory for text. International journal of educational research, 1997, 27(3): 255-265.

[26] Voss JF, Wiley J. Developing understanding while writing essays in history. Journal of educational psychology, 1999, 91(2): 301-311.

[27] Hagen ÅM, Bråten I, Brøten I. Relationships between spontaneous note-taking, self-reported strategies and comprehension when reading multiple texts in different task conditions. Journal of research in reading, 2014, 37(S1): S141-S157.

[28] Capra R, Marchionini G, Velasco-Martin J, et al. Tools-at-hand and learning in multi-session, collaborative search. Proceedings of the SIGCHI conference on human factors in computing systems. New York: ACM, 2010: 951-960.

[29] Song X, Liu C, Liu H. Characterizing and exploring users' task completion process at different stages in learning-related tasks. Proceedings of the Association for Information Science and Technology, 2018, 55(1): 460-469.

[30] Wu JY, Xie C. Using time pressure and note-taking to prevent digital distraction behavior and enhance online search performance: perspectives from the load theory of attention and cognitive control. Computers in human behavior, 2018, 88: 244-254.

[31] Wu JY. The predictive validities of individual working-memory capacity profiles and note-taking strategies on online search performance. Journal of computer assisted learning, 2020, 36(6): 876-889.

[32] Liu Chang, Song Xiaoxuan, Yang Ziao. Exploration of user learning behavior and process in information search. Journal of Academic Libraries, 2019, 37(4): 36-45.

[33] Roy N, Torre M V, Gadiraju U, et al. Note the highlight: incorporating active reading tools in a search as learning environment. Proceedings of the 2021 conference on human information interaction and retrieval. New York: ACM, 2021: 229-238.

[34] Song Xiaoxuan, Liu Chang. Assessment of user knowledge level before and after searching and analysis of its changes. Library and Information Service, 2018, 62(2): 108-116.

[35] Peterson ER, Deary IJ, Austin EJ. Are intelligence and personality related to verbal-imagery and wholist-analytic cognitive styles? Personality & individual differences, 2005, 39(1): 201-213.

[36] Song Xiaoxuan, Liu Chang. Research on user information source selection and use strategies in learning-related search. Journal of the China Society for Scientific and Technical Information, 2019, 38(6): 655-666.

[37] Rieh SY, Collins-Thompson K, Hansen P, et al. Towards searching as a learning process: a review of current perspectives and future directions. Journal of information science, 2016, 42(1): 19-34.

[38] Riding R, Cheema I. Cognitive Styles—an overview and integration. Educational psychology: an international journal of experimental educational psychology, 1991, 11(3/4): 193-215.

[39] Liu Hanrui, Liu Chang. Impact of cognitive style and topic familiarity on search interaction behavior in learning-related tasks. Information Studies: Theory & Application, 2018, 41(4): 56-62.

[40] Mueller PA, Oppenheimer DM. The pen is mightier than the keyboard: advantages of longhand over laptop note taking. Psychological science, 2014, 25(6): 1159-1168.

[41] Craik FIM, Lockhart RS. Levels of processing: a framework for memory research. Journal of verbal learning and verbal behavior, 1972, 11(6): 671-684.

[42] Kiewra KA. Investigating notetaking and review: a depth of processing alternative. Educational psychologist, 1985, 20(1): 23-32.

[43] Cheng Bo, Huang Xiaobin. Research on current application status of web content analysis at home and abroad. Library and Information Service, 2007(9): 45-50.

[44] Sun Xiaojun, Zhou Zongkui. Exploratory factor analysis and its main problems in application. Psychological Science, 2005, 28(6): 1440-1442.

[45] Li Can, Xin Ling. Research on evaluation methods for reliability and validity of questionnaires. Chinese Journal of Health Statistics, 2008, 25(5): 541-544.

[46] Crescenzi A, Li Y, Zhang Y, et al. Towards better support for exploratory search through investigation of notes-to-self and notes-to-share. Proceedings of the 42nd international ACM SIGIR conference on research and development in information retrieval. New York: ACM, 2019: 1093-1096.

[47] Pepper RJ, Mayer RE. Notetaking as a generative activity. Journal of educational psychology, 1978, 78(4): 514-522.

[48] Fisher M, Goddu MK, Keil FC. Searching for explanations: how the internet inflates estimates of internal knowledge. Journal of experimental psychology: general, 2015, 144(3): 674-687.

Author Contributions:
Meng Gaohui: Literature review, data collection, analysis, and initial draft
Song Xiaoxuan: Data collection, manuscript revision
Zhang Xiaoyue: Participated in data analysis
Liu Chang: Research design guidance, manuscript revision and finalization

Note: Figure translations are in progress. See original paper for figures.