

Postprint: Research on Retrieval Behavior for Perceived Task Difficulty

Authors: Huang Kun, Chen Jiaqi, Liu Chang

Date: 2023-04-01T15:51:23+00:00

Abstract

[Purpose/Significance] In the context of learning-oriented search, this study explores interaction behavior indicators that exert stable and significant influence on reflecting users' perceived task difficulty. [Method/Process] Behavioral experiments were conducted within classroom teaching settings, with students from the "Information Technology Application" course serving as research subjects. Five search tasks were assigned and processes were recorded using Morae software, yielding 170 valid retrieval behavior data entries from 39 students. Based on screen recording data and task difficulty perception evaluation data, behavior indicators exhibiting significant differences across difficulty levels were analyzed. Combined with two experimental datasets from 2016 and 2017 conducted by the research group, and through horizontal comparison with existing related research findings, the stability of behavior indicators showing significant differences under varying task difficulties was summarized. [Results/Conclusions] A stable behavior indicator system comprising 26 indicators was ultimately obtained. Behavioral characteristics of users under difficult tasks were described from the perspectives of exploratory search process and behavior types, and directions for future research were identified.

Full Text

Preamble

Research on Search Behaviors Oriented to Task Difficulty Perception

Huang Kun¹, **Chen Jiaqi**¹, **Liu Chang**²

¹School of Government, Beijing Normal University, Beijing 100875

²Department of Information Management, Peking University, Beijing 100871

Abstract: [Purpose/Significance] Against the backdrop of learning-related search, this study explores interactive behavior indicators that have stable and significant influence on reflecting users' perceived task difficulty.

[Method/Process] Relying on classroom teaching, behavioral experiments were conducted with students from the “Information Technology Application” course as research subjects. Five search tasks were assigned, and Morae software was used for process recording. A total of 170 valid retrieval behavior records from 39 students were collected. Based on screen recording data and task difficulty perception evaluation data, behavioral indicators showing significant differences under varying difficulty levels were analyzed. Combined with data from two experimental studies conducted by the research team in 2016 and 2017, as well as findings from existing related research, horizontal comparisons were made to summarize the stability of behavioral indicators that showed significant differences across different task difficulty levels. [Result/Conclusion] A stable behavioral indicator system with 26 indicators was ultimately obtained. The characteristics of user behavior under difficult tasks are described from the perspective of exploratory search processes and behavior types, and directions for future research are identified.

Keywords: task difficulty; retrieval behavior; interactive information retrieval; experimental study

Classification Number: G252.7

DOI: 10.13266/j.issn.0252-3116.2022.03.010

Exploring the connections between multi-dimensional attribute characteristics of tasks and observable features of search interaction processes has become a research hotspot in the field of interactive information retrieval in recent years [1]. Task difficulty is one of the important attributes of tasks [2-3], influencing retrieval interaction behavior, retrieval performance, and retrieval experience [4]. In learning-related search, task difficulty also affects learning outcomes and learning experience during the search process [5]. Difficulty generally refers to the level of difficulty or ease of an item (such as a task or question) [6]. Users’ perception of search task difficulty may stem from various difficulties in the search process and is related to users, tasks, and the interaction process [7], including difficulties in understanding information needs, conducting search processes, and interpreting result relevance [8]. To improve users’ search performance, search experience, and learning gains, researchers have explored the relationship between task difficulty attributes and retrieval behavior, aiming to construct task difficulty prediction models based on retrieval behavior to provide references for retrieval systems to timely understand users’ search intentions and dilemmas and offer corresponding interventions and support [9-12].

In terms of information search task difficulty, many researchers have conducted a series of studies. Previous work by this research team has carried out a systematic review [13]. It can be seen that many studies have explored the characteristics and patterns of retrieval behavior under different task difficulty levels from perspectives such as reasons for task difficulty, web information retrieval, and retrieval of resources with specific topics or media types, as well as task difficulty prediction based on behavior. These studies mostly selected behavioral indicators from aspects such as query, browsing, click, and save, with some re-

search also incorporating retrieval strategies [14], search paths [15], and other features. Some researchers have comprehensively considered behavioral characteristics from a temporal stage perspective [11]. These studies often collected data through laboratory research and conducted correlation analysis between behavioral data and difficulty evaluation data. However, findings from different studies are not entirely consistent. For example, in two studies by C. Liu et al. [11-12], it was found that users viewed more content pages in simple tasks than in difficult tasks; whereas in studies by J. Arguello [9], J. Gwizdka and I. Spence [15], and J. Kim [16], it was found that users viewed more web pages in difficult tasks. Which retrieval behavior indicators can better distinguish user states under different task difficulty perceptions? This affects the reliability and stability of subsequent task difficulty prediction based on behavior. Currently, few researchers have conducted comparative work across multiple experiments. Therefore, this study combines three different information retrieval behavior experiments and conducts comprehensive horizontal comparisons with findings from similar studies to identify more stable behavioral indicators that have significant influence on task difficulty perception, providing references for future prediction work.

2 Research Design

2.1 Research Questions

This study uses learning-related search as the background, combined with actual course teaching, to conduct retrieval behavior experiments. By collecting students' behavioral data and task difficulty perception data when completing search tasks, it analyzes behavioral indicators that significantly influence difficulty perception and compares findings with different experiments to evaluate indicator stability. Specific research questions include: (1) Do retrieval behavior characteristics of query, click, mouse, view, and time categories, as well as search strategy characteristics, show significant differences in search processes with different task difficulty perceptions? (2) Among the significant indicators found in (1), which behavioral indicators have better stability in reflecting task difficulty perception across different experimental studies?

2.2 Retrieval Behavior Indicators

Based on existing research [9-11], this study established five categories of retrieval behavior: query, click, mouse, view, and time. Combined with retrieval strategy types proposed by R. Fidel [17], these were divided into operational and conceptual aspects. Operational strategies refer to operations that adjust retrieval result sets through retrieval system functions, while conceptual strategies refer to operations that adjust retrieval result sets by modifying the conceptual meaning of retrieval queries. Additionally, web pages browsed by users were categorized into three types: "login page" before entering search terms, "SERP page" displaying retrieval result lists, and "Doc page" displaying summaries or full text of single retrieval results. Therefore, this study constructed a

retrieval behavior indicator system with 7 feature categories and 50 behavioral indicators: query features (10 items), click features (6 items), mouse features (9 items), view features (6 items), time features (14 items), operational strategies (1 item), and conceptual strategies (4 items). Specific behavioral indicators are shown in Table 1 .

2.3 Task Design

This study recruited undergraduate students from the university-wide general education course “Information Technology Application,” which mainly teaches Access database and VBA programming-related knowledge to lower-level undergraduates. The research team (hereinafter referred to as “the team”) collected questions commonly concerned by students in previous years, combined with task designs from existing research, and drew on educational objective classification [18] and characteristics of cognitive complexity in interactive retrieval tasks [19] to design four categories of learning-related search tasks: understanding, analysis, evaluation, and creation, with five specific tasks. Task descriptions are shown in Table 2 .

2.4 Data Collection

This experiment (the third experiment) was conducted in the spring 2020 semester of the “Information Technology Application” course. In the first class, tasks were distributed as homework assignments. Students used Morae software to record the process of information retrieval and homework writing. Before completing the assignments, a 5-point Likert scale was used to collect students’ perceptions of task difficulty, familiarity, and interest. After completing each retrieval task, a 5-point Likert scale was also used to collect students’ perceptions of task difficulty, reasons for difficulty, and retrieval performance (including relevance, usefulness, confidence level, and sense of success). A total of 39 students submitted assignments, including 10 males and 29 females. Among them, 36 were from the School of Government, 2 from the School of Economics and Business Management, and 1 from the Law School. Additionally, 36 were freshmen, 2 were sophomores, and 1 was a junior. All students used personal computers to record their search processes with Morae. Due to Morae software lagging and restarting, some task process recordings were incomplete. The actual collection yielded 170 valid retrieval behavior records.

Because of Morae recording frequency issues, six behavioral indicators had certain missing data: SERP page mouse scrolling down behavior, SERP page mouse scrolling up behavior, Doc page mouse scrolling down behavior, Doc page mouse scrolling up behavior, mouse movement distance, and total document editing duration. However, the highest missing rate for behavioral variables in the experiment was 32.35%, which is below the missing value threshold [20]. Therefore, based on comparative analysis of various missing data processing methods [21], this study selected the EM algorithm proposed by A.P. Dempster

et al. [22] to fill in missing behavioral variables caused by recording software issues.

Furthermore, this study comprehensively utilized two previous experimental studies by the team: (1) Experiment One [23] was conducted in 2016, designing three categories of five tasks (factual, exploratory, and abstract) based on task complexity [24], and randomly recruited 30 students for laboratory research. The subjects included 5 males and 25 females, with 8 undergraduates and 22 graduate students from 6 schools including the School of Government and the School of Geography. (2) Experiment Two [25] was conducted in 2017, designing five categories of five tasks (memory, understanding, analysis, evaluation, and creation) based on the educational objective classification method [18], and randomly recruited 30 students for laboratory research. This experiment included 9 males and 21 females, with 16 undergraduates and 14 graduate students from 21 schools. Both experiments collected users' evaluations of task difficulty perception before and after task completion, as well as behavioral data from the search process. This study uses data from three experiments and conducts horizontal comparative analysis with findings from similar studies to identify more stable behavioral indicators that influence task difficulty perception across different experimental tasks and subject compositions. For concise expression, "Experiment One" refers to the 2016 study, "Experiment Two" refers to the 2017 study, and "Experiment Three" refers to the latest study conducted in this paper.

3 Data Analysis and Discussion

3.1 Basic Task Difficulty Perception

Using users' post-retrieval ratings of task difficulty as the basis for judging perceived task difficulty, the overall difficulty levels across the three experiments were comparable. The distribution of task difficulty and overall mean and standard deviation for the three experiments are shown in Table 3 .

3.2 Analysis of Significant Behavioral Indicators

In the three experiments, post-retrieval task difficulty evaluations were selected. Ratings of 4 and 5 were classified as difficult tasks, 3 as moderately difficult, and 1 and 2 as easy tasks. All three experimental datasets used the Kruskal-Wallis H test method to compare differences and significance in users' retrieval behaviors across different difficulty tasks. Table 4 lists the values of behavioral indicators showing significant differences under different task difficulty levels across the three experiments.

Based on analysis of the three experimental datasets: (1) Among 6 query indicators, 4 showed significant differences across different difficulty tasks in all three experiments, all demonstrating significantly higher behavior counts under high-difficulty tasks compared to low-difficulty tasks. (2) Among 5 click indicators, 2 showed significant differences, also demonstrating significantly higher

behavior counts under high-difficulty tasks. However, behavioral differences between moderately difficult and highly difficult tasks were not always significant, though differences between easy tasks and non-easy tasks were significant. (3) Among 7 mouse indicators, behavioral data from the first two experiments were incomplete. Comparisons between Experiment One and Three, and between Experiment Two and Three, showed that although these indicators demonstrated differences across task difficulty levels, their patterns were not entirely consistent. For example, moderately difficult tasks showed significantly higher “Doc page mouse scrolling down behavior” than easy and difficult tasks, not following a consistent increasing pattern from easy to difficult. (4) Among 5 view indicators, all showed significant differences across different difficulty tasks in all three experiments, with significantly lower behavior counts under easy tasks compared to moderate and difficult tasks. (5) Among 7 time indicators, 3 showed significant differences across different difficulty tasks, all demonstrating significantly longer durations under difficult tasks compared to easy tasks. (6) Among 4 retrieval strategy indicators, 2 showed significant differences across different difficulty tasks in all three experiments, also demonstrating significantly higher usage under difficult tasks.

Indicators showing significant differences across different task difficulty perceptions in all three experiments are termed stable indicators, while others are classified as unstable indicators and eliminated. This resulted in 26 stable behavioral indicators. These indicators signify characteristics that reflect users’ difficulty perception differences across different experimental task contexts. Additionally, this study conducted horizontal comparisons between the 26 indicators and existing research literature and findings. Indicators with consistent influence directions of task difficulty perception on retrieval behavior are called common indicators, inconsistent ones are called difference indicators, and those not appearing in existing research are called new indicators.

To analyze the relationship between retrieval behavior indicators and exploratory search processes, this study draws on the two basic processes of browsing and focused searching proposed by R.W. White et al. [26], categorizing retrieval behaviors into exploratory browsing and focused searching. Exploratory browsing mainly includes discovering, learning, and investigating, while focused searching mainly includes query reformulation and result analysis. As shown in Table 5, the 26 retrieval behavior indicators are categorized into initial retrieval and retrieval process stages, with the retrieval process further divided into focused searching and exploratory browsing.

3.3 Discussion

As shown in Table 5:

- (1) From the perspective of users’ interpretation of tasks and conception of retrieval strategies at the initial stage, users spent significantly longer time on the retrieval login page (the duration from opening the retrieval inter-

face to entering search terms and clicking search) under difficult tasks, reflecting that users need more time to contemplate and think about retrieval strategies. C.C. Kuhlthau [33] also noted in the ISP model that at the beginning of task execution, users have a vague sense of uncertainty about retrieval goals. When users feel they may lack required domain knowledge or are not proficient, they perceive greater task difficulty and exert more effort to solve the problem [34]. Therefore, they may spend longer time on understanding task descriptions, connecting existing domain knowledge, selecting keywords, constructing search queries, selecting information sources, and anticipating retrieval results. Regarding the “total login page duration” indicator, similar indicators exist in previous research but are not identical. This study defines “login page” as the retrieval interface before entering search terms, while J. Arguello [9] defined “loading page” as the time to open pages after clicking retrieval results, more similar to part of Doc page browsing. Therefore, this is considered a new indicator.

- (2) From the perspective of task retrieval execution, users continuously conduct retrieval attempts and make selections while viewing, analyzing, and comparing retrieval results. Multiple studies have found behavioral indicators with significant differences between difficult and easy tasks in this process.

First, in focused searching, 13 behavioral indicators significantly distinguish different task difficulty levels, including 9 common indicators and 4 new indicators. When facing more difficult tasks, users’ query and click behaviors are significantly more frequent than under easy tasks. Under difficult tasks, users continuously adjust and reformulate search queries, use more different search terms, and show significantly higher result click counts, page-turning counts, and click percentages during queries compared to easy tasks. This indicates that users view more information and click more frequently during queries. Similarly, A. Inthiran et al. [28], J. Arguello [9], J. Liu et al. [11], A. Aula [27], and O. Dan [29] all found that users make significantly more queries in difficult tasks than in easy tasks. J. Arguello [9], J. Liu et al. [11], A. Aula [27] showed that query length and number of query terms are higher in difficult tasks. Moreover, users demonstrate more retrieval adjustments in difficult tasks. J. Kim’s research also found that users’ retrieval adjustment frequency significantly increases in difficult tasks [16]. As task complexity increases, the number of retrieval adjustment operations also increases [14].

In addition to specific retrieval behavior indicators, this study added 4 indicators from retrieval strategies, all of which can distinguish task difficulty. In operational strategies, difficult tasks mainly manifest as users employing more diverse information source types, including different platforms and different functional areas within the same platform. For conceptual strategies, when handling difficult tasks, users more carefully extract retrieval terms from task descriptions and are inspired during the search process to propose more meaningful query terms

not mentioned in the task description. These new query terms may come from users' learning or divergent thinking about results, or from system-recommended resources. X. Niu and D. Kelly [35] explored users' use of retrieval systems with query suggestion functions, finding that participants incorporated query suggestions very quickly, and users with less professional knowledge and experience used more query suggestions and saved more documents. Participants also used more query suggestions in more difficult tasks.

Second, in exploratory browsing, 12 behavioral indicators significantly distinguish different task difficulty levels, including 8 common indicators, 3 new indicators, and 1 difference indicator. Common indicators show that when completing difficult tasks, users' maximum depth of result viewing, mouse hover counts on potentially useful information, view counts of all SERP pages and unique SERP pages, and total web page counts and unique web page counts are all significantly higher than in easy tasks. Additionally, SERP dwell time and total task duration are significantly longer in difficult tasks. These behaviors are basic activities for users to study, filter, and judge result relevance. Evidently, longer reading time on result pages and viewing more different results may indicate users' difficulty perception at that time. Similarly, J. Arguello [9], X. Hu and N. Kando [31], and L. Fraser and C. Locatis [30] all found that users' mouse click counts significantly increase in difficult tasks. J. Arguello [9] found that mouse and scrollbar operations occur more frequently in difficult tasks. J. Arguello [9], J. Liu et al. [11], X. Hu and N. Kando [31], D. Hienert et al. [32], J. Gwizdka and I. Spence [15], A. Inthiran et al. [28], and J. Kim [16] all found that users spend significantly more time on difficult tasks than easy tasks.

Based on common indicators, this study added 3 new indicators. The "average mouse hover ranking" indicator was optimized based on previous research. Its significant results indicate that in difficult tasks, users conduct deeper browsing of result lists rather than just browsing top-ranked results. Additionally, this study distinguished between SERP page and Doc page dwell times, comprehensively calculating "total web page dwell time." The significant results of this indicator also reflect users' cognitive effort in browsing, understanding, and learning search content. Meanwhile, "total document editing duration" is a new indicator specific to the learning-related search conducted in this study, reflecting users' cognitive effort in information integration and summarization to complete task objectives, also showing that users spend significantly longer time and effort on difficult tasks than easy tasks.

In exploratory browsing, there is also 1 difference indicator: "total number of unique Doc pages viewed." J. Liu et al. [11] found that "total number of unique Doc pages viewed" showed a decreasing trend with increasing task difficulty. Comparing the task designs of the two experiments, the experiments conducted by this team either required users to answer questions or complete research report writing, which created higher demand for viewing Doc pages. In contrast, the studies by J. Arguello [9] and J. Liu et al. [11] only required users to save or mark relevant results, which may be the reason for the experimental result

differences.

Based on the above discussion, through multiple experimental comparisons and horizontal comparisons with similar studies, this study summarizes behavioral indicators with significant effects on task difficulty perception judgment under different research conditions, helping to reveal more stable relationships between task difficulty perception and retrieval behavior. This lays a foundation for further quantitative mapping and research, and provides references for behavioral feature selection and weight design in future task attribute prediction research. Simultaneously, combined with analysis of typical exploratory search processes, theoretical support is provided for explaining how these behaviors reflect task difficulty perception.

This study integrated literature review and previous research foundations, conducted experimental research relying on classroom teaching, used students from the “Information Technology Application” course as research subjects, and utilized Morae software to conduct remote retrieval behavior experiments, collecting 170 valid retrieval behavior records from 39 students. Based on screen recording data and task difficulty perception evaluation data, behavioral indicators showing significant differences under different difficulty levels were analyzed. Combined with experimental results from studies conducted by the team in 2016 and 2017, as well as comprehensive comparisons with existing related research findings, 26 behavioral indicators with significant influence on distinguishing task difficulty perception across different studies were summarized.

In social science experimental research, repeating and reproducing identical experimental environments and conditions faces many difficulties and challenges. For example, participants recruited for experiments have subjectivity and individual differences, and task designs used as experimental stimuli have difficulties in precisely measuring and controlling the degree of difficulty that triggers participants’ difficulty perception. However, through comparisons of three experimental results, it can be seen that using a five-level scale to measure subjective task difficulty perception, conducting experiments with unified laboratory machines versus personal computers, can still reveal highly stable retrieval behavior indicators under different experimental task designs. These indicators show significant differences across different task difficulty perceptions, laying a foundation for more stable task difficulty perception prediction. Future research can also more deeply explore factors affecting indicator stability, different feature combination methods, or combine temporal dimensions to explore modeling strategies and methods for difficulty prediction based on retrieval behavior. Moreover, future research can combine specific retrieval context characteristics to establish quantitative mapping relationships between users’ task difficulty perception and multi-level, different granularities of retrieval behavior, achieving task difficulty identification and judgment functions based on behavior to embed in information systems. This would enable retrieval systems to not only capture users’ potential search difficulties and provide timely interventions and adjustments through implicit user observation but also optimize task assignment

mechanisms, allowing users to continue and complete search tasks as much as possible within their capabilities, avoiding withdrawal or cessation due to task difficulty.

The main limitation of this study is that although multiple rounds of experiments were conducted and comparisons were made with similar studies, revealing relatively stable relationships between task difficulty perception and retrieval behavior across different studies, the overall scale of experimental subjects was relatively limited, which has certain limitations in overall sample representativeness and requires testing in larger-scale search contexts. Additionally, most laboratory research task designs come from experimenter assignments, which are somewhat distant from users' real tasks. Therefore, subsequent log data comparisons and validations could better test the universality of the behavioral indicators extracted in this study for judging users' task difficulty perception. Furthermore, existing indicators are all cumulative results of retrieval processes, providing insufficient real-time reflection of task difficulty perception during retrieval processes, and differences in the degree of difficulty perception distinguished by different indicators are not considered. Future research can conduct more fine-grained studies on task difficulty perception and its changes during search processes, and explore patterns related to retrieval behavior more deeply. Simultaneously, richer observable data such as eye-tracking, facial expressions, and logs can be combined to provide support for task difficulty analysis, promoting information retrieval systems' capture of users' search cognition and emotional states, and improving users' retrieval performance and experience.

References

- [1] LIU J Q. Deconstructing search tasks in interactive information retrieval: a systematic review of task dimensions and predictors [EB/OL]. [2021-10-29]. <https://doi.org/10.1016/j.ipm.2021.102522>. [2] LI Y L, BELKIN N J. A faceted approach to conceptualizing tasks in information seeking [J]. *Information processing & management*, 2008, 44(6): 1822-1837. [3] KIM S, SOERGEL D. Selecting and measuring task characteristics as independent variables [EB/OL]. [2021-10-29]. <https://doi.org/10.1002/meet.14504201111>. [4] CRESCENZI A, CAPRA R, ARGUELLO J. Time pressure, user satisfaction and task difficulty [J]. *Proceedings of the American Society for Information Science and Technology*, 2014, 50(1): 1-4. [5] Liu Chang, Song Xiaoxuan, Yang Ziao. Exploration of user learning behavior and process in information search [J]. *Journal of Academic Libraries*, 2019, 37(4): 36-45. [6] Zhu Peiliang. *Psychometrics: Theory and Application* [M]. Hefei: University of Science and Technology of China Press, 2008. [7] LIU J J, LI Y, HASTINGS S K. Simplified scheme of search task difficulty reasons [J]. *Journal of the Association for Information Science and Technology*, 2019, 70(5): 526-529. [8] BELL D J, RUTHVEN I. Searcher's assessments of task complexity for Web searching [C]//*Proceedings of the 26th European conference on information retrieval (ECIR 2004)*. Berlin: Springer, 2004: 57-71. [9] ARGUELLO J. Predicting search task difficulty [C]//*Proceedings of*

European conference on information retrieval. Cham: Springer, 2014: 88-99. [10] LIU J J, GWIZDZKA J, LIU C, et al. Predicting task difficulty for different task types [C]//Proceedings of the American Society for Information Science and Technology, 2010, 47(1): 1-10. [11] LIU J J, LIU C, COLE M, et al. Exploring and predicting search task difficulty [C]//Proceedings of the 21st ACM international conference on information and knowledge management. Maui: ACM, 2012: 1313-1322. [12] LIU C, LIU J J, BELKIN N J. Predicting search task difficulty at different search stages [C]//Proceedings of the 23rd ACM international conference on conference on information and knowledge management. Shanghai: ACM, 2014: 569-578. [13] Huang Kun, Chen Jiaqi, Zheng Mingxuan, et al. Review of information search task difficulty research [J]. Journal of Information Resources Management, 2020, 10(4): 88-98. [14] WILDERMUTH B M, KELLY D, BOETTCHER E, et al. Examining the impact of domain and cognitive complexity on query formulation and reformulation [J]. Proceedings of the American Society for Information Science and Technology, 2006, 43(1): 1-22. [15] GWIZDZKA J, SPENCE I. What can searching behavior tell us about the difficulty of information tasks? a study of Web navigation [C]//Proceedings of CHI'06 extended abstracts on human factors in computing systems. Montréal: ACM, 2006: 959-964. [16] KIM J. Task difficulty as a predictor and indicator of Web searching interaction [C]//Proceedings of the 3rd ACM SIGIR conference on human information interaction and retrieval (CHIIR). New Brunswick: ACM, 2018: 72-81. [17] FIDEL R. Moves in online searching [J]. Online review, 1985, 9(1): 61-74. [18] ANDERSON L W, KRATHWOHL D R, AIRASIAN P W, et al. Taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives [J]. European legacy, 2001, 114(458): 1013-1014. [19] KELLY D, ARGUELLO J, EDWARDS A, et al. Development and evaluation of search tasks for IIR experiments using a cognitive complexity framework [C]//Proceedings of the 2015 international conference on the theory of information retrieval. Northampton: ACM, 2015: 101-110. [20] BARZI F, WOODWARD M. Imputations of missing values in practice: results from imputations of serum cholesterol in 28 cohort studies [J]. American journal of epidemiology, 2004, 160(1): 34-45. [21] Pang Xinsheng. Comparison of missing data processing methods [J]. Statistics & Decision, 2010(24): 152-155. [22] DEMPSTER A P, LAIRD N M, RUBIN D B. Maximum likelihood from incomplete data via the EM algorithm [J]. Journal of the Royal Statistical Society, series B (Methodological), 1977, 39(1): 1-38. [23] Zhang Lulu, Huang Kun. Research on digital library user information retrieval behavior based on cognitive style [J]. Journal of the China Society for Scientific and Technical Information, 2018, 37(11): 1164-1174. [24] BYSTRÖM K, JÄRVELIN K. Task complexity affects information seeking and use [J]. Information processing & management, 1995, 31(2): 191-213. [25] Wang Kaifei, Huang Kun. Research on influencing factors of digital library user retrieval failure [J]. Library and Information Service, 2019, 63(11): 25-34. [26] WHITE R W, ROTH R A. Exploratory search: beyond the query-response paradigm [J]. Synthesis lectures on information concepts, retrieval, and services, 2009, 1(1): 1-98. [27] AULA A, KHAN R M, GUAN Z. How does search behavior change as search becomes more

difficult? [C]//Proceedings of the SIGCHI conference on human factors in computing systems. Atlanta: ACM, 2010: 35-44. [28] INTHIRAN A, ALHASHMI S M, AHMED P K. A user study on the information search behaviour of medical students [J]. Malaysian journal of library & information science, 2015, 20(1): 61-73. [29] DAN O, DAVISON B D. Measuring and predicting search engine users' satisfaction [J]. ACM computing surveys, 2016, 49(1): 1-34. [30] FRASER L, LOCATIS C. Effects of link annotations on search performance in layered and unlayered hierarchically organized information spaces [J]. Journal of the American Society for Information Science and Technology, 2001, 52(14): 1255-1261. [31] HU X, KANDO N. Task complexity and difficulty in music information retrieval [J]. Journal of the Association for Information Science and Technology, 2017, 68(7): 1711-1723. [32] HIENERt D, MITSUI M, MAYR P, et al. The role of the task topic in Web search of different task types [C]//Proceedings of 3rd ACM SIGIR conference on human information interaction and retrieval (CHIIR). New Brunswick: ACM, 2018: 72-81. [33] KUHLTHAU C C. Seeking meaning: a process approach to library and information services [M]. New York: Libraries Unlimited Inc, 2004. [34] DERVIN B. Sense-making theory and practice: an overview of user interests in knowledge seeking and use [J]. Journal of knowledge management, 1998, 2(2): 36-46. [35] NIU X, KELLY D. The use of query suggestions during information search [J]. Information processing & management, 2014, 50(1): 218-234.

Author Contributions:

Huang Kun: Determined research topic, proposed research design, overall manuscript revision;

Chen Jiaqi: Participated in research design, conducted data collection and analysis, wrote manuscript;

Liu Chang: Participated in research design and data discussion analysis.

Research on the Search Behaviors Oriented to Task Difficulty Perception

Huang Kun¹, Chen Jiaqi¹, Liu Chang²

¹School of Government, Beijing Normal University, Beijing 100875

²Department of Information Management, Peking University, Beijing 100871

Abstract: [Purpose/significance] This study uses learning-related search as the background to explore interactive behavior indicators that have stable and significant impact on users' perception of task difficulty. [Method/process] Relying on classroom teaching, behavioral experiments were conducted with students from the "Information Technology Application" course as research subjects. Five search tasks were assigned, and Morae software was used to record the process remotely. A total of 170 valid retrieval behavior records from 39 students were collected. Based on screen recording data and task difficulty perception evaluation data, behavioral indicators showing significant differences under different difficulty levels were analyzed. Combined with experimental data from two stud-

ies conducted by the research team in 2016 and 2017, as well as findings from existing related research, horizontal comparisons were made to summarize the stability of behavioral indicators showing significant differences across different task difficulty levels. [Result/conclusion] Finally, a stable behavioral indicator system with 26 indicators was obtained. The characteristics of user behavior in difficult tasks are described from the perspective of exploratory search processes and behavior types, and future research directions are identified.

Keywords: task difficulty; search behavior; interactive information retrieval; experimental study

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

Source: ChinaXiv — Machine translation. Verify with original.