Process-learning as a factor in evaluation

Catherine L. Smith
Kent State University
School of Library and Information Science
330-672-2116
csmit141@kent.edu

ABSTRACT

This position paper argues that process-learning is a key factor in the evaluation of interactive search systems. Process-learning is defined as an outcome of search interaction, whereby the searcher accumulates habits and strategies that guide interaction. Given that local system performance (at the query level) is dependent on the searcher's actions (e.g. query terms) and cognitions (e.g. judgments on results), a system is likely to reinforce the search actions that produce successful results most often. This is a desirable outcome for needs that are well-matched with the system's optimal performance range (the types of needs for which it is tuned); however, it often leaves searchers with few alternative processes (skills) for other types of needs. We speculate that when searchers are assisted in learning alternative approaches, behavior will be more predictable across a broader range of needs, thereby tightening the coupling of feedback between the system and its user. We argue that where evaluation paradigms include the searcher's process-learning, designs will benefit from deeper theories for user-centered performance evaluation and optimization.

CCS Concepts

• **Information retrieval ➝ Users and interactive retrieval ➝ Evaluation of retrieval results**

Keywords

Search interaction; performance evaluation; learning.

1. INTRODUCTION

This position paper proposes the explicit inclusion of the user's learning in the evaluation framework put forth by Järvelin, et al. (2015). We start with a brief summary of the framework, with emphasis on the process elements associated with search interaction. We then discuss the role of learning in a programmatic view of interaction, and two forms of learning associated with interactive processes. Next, an argument is put forth on the value of system designs that support searchers in gaining a broader repertoire of effective search actions. We then speculate on the ways in which performance evaluation and ultimately system performance, may be enhanced by designs focused on searchers' process-learning. The brief closing reviews how process-learning is integral to user-centered evaluation.

1.1 The process view of evaluation.

Vakkari and Kekalainen [15] and Järvelin, et al. [5] have proposed the application of program theory to problems of evaluation in information interaction, most recently in a multifactorial evaluation framework for what is termed *task-based information interaction* (TBII). In its broad view of information needs and problems, the framework conceptualizes a nested set of information *processes*, each of which contains internal sub-processes. Processes are defined as actions and cognitions that transform *inputs* (independent variables) into *outputs* (dependent variables), which in turn affect *outcomes*. Outcomes are the synthesized products and benefits of outputs. The framework provides a basis for discussion on the components of interactive information processes, and most importantly, on their interdependencies. The goal is a structure in which researchers may position the objectives, methods, and results of evaluation research. The framework motivates demarcation among various goals and forms of evaluation, and importantly, their attendant standards for design, analysis, and reporting methods. The approach illuminates the basis for tradeoffs made in internal and external validity, with the possibility of tighter linkages between results flowing from different evaluation approaches. The ultimate goal is an understanding of *how* and *why* design alternatives affect the quality and value of information outcomes.

In presenting the framework, the authors apply it to learning-tasks as an example. Included among the processes described are *information searching* and *item selection*, which are connected by *search results*, an output of searching. The sub-processes listed parallel those of Marchionini's [8] *plan and execute* sub-processes and include: selection of the system/collection; query formulation; item examination; item judgment; and the selection of items for use. The synthesized outcomes of these processes are (1) information that serves the current search task and (2) knowledge of other available information. This view focuses on change in the searcher's knowledge state with respect to the current task, with no account for the searchers learning with respect to the process. In discussing the limitations of the model, the authors acknowledge that the searcher "cannot avoid learning about …. the process" (3:24, [5]). We argue here that this learning, which we call *process-learning*, is a key outcome of and effect upon search interaction. Further, we argue that where an evaluation paradigm takes process-learning into account explicitly, the interdependent effects of system performance and search behavior may be better understood. Next we discuss two forms of process-learning.

1.2 Two forms of process-learning.

One form of learning, *proceduralization* [1], is a byproduct of the repeated use of a process, which results in the formation of habits for action and cognition. For example, searchers have a propensity to scan and click at the top of results lists, even when items below are more relevant (*trust bias*, see [6]). Habits of this type are generally deployed automatically, with little cognitive effort, thus freeing attention for other matters. Habits like trust-bias are likely to be efficient when using a familiar system for needs that match well with the system; however, they are less likely to be optimal for poorly matched needs and goals, or for use of an unfamiliar system. In these cases, the searcher may be best served by the use of alternative actions.

Before a searcher can use alternative actions, they must be learned. Generally, learning is most efficient when it starts with explicit (declarative) knowledge of a plan or schema for the action, and when this is scaffolded for the learner [16]. Initial use of a new action is cognitively demanding, as it requires explicit attention to the learned process itself. For search, this additional cognitive load is required for processes that are likely ancillary to the main task of information seeking and sense-making, thus, learning and maintaining alternative actions requires support.

1.3 Learning alternative actions.

Because current systems provide little support for learning alternative actions (or because alternatives may have little value), most searchers rely on the simple processes that work for most needs [9]. In the face of poor local system performance (need-specific performance at the query level), searchers rely habitually on repeated use of the *vary* tactic (switching query terms) [14]. Adaptive change in behavior may simply be more rapid iteration of this one tactic [12]. Some searchers develop their own creative solutions for more effective action (e.g., opening multiple browser windows for search on multi-faceted needs) [17], but struggling searchers know few options for improving the performance of the system. Affordances such as query suggestion provide local alternatives for action (need-specific alternatives at the query level: "select these words"). These mechanisms are designed with little consideration of their reinforcement of the searcher's *localized thinking* (e.g., need-specific attention to the terms readily available in memory [11]). Interventions and affordances for reflection on search processes [3], delivered at the appropriate time [7] along with guidance [4], may scaffold deeper process-learning for more cognitively demanding yet more effective actions. When evaluation is focused on local effects to the exclusion of process-learning, the community forgoes opportunities to enhance local performance by affecting the searcher's global performance (the effective use of a repertoire of alternative actions over the long-term).

1.4 Learning and performance objectives.

It is well established that system performance is affected greatly by differences in the behavior of searchers [13]. Thus, when real searchers are used in evaluation, performance is dependent on the characteristics of searchers (e.g., cognitive capacities, personal traits, habits, prior knowledge) and their actions (query terms, relevance judgments). In global evaluation that measures average performance, this variability is considered a nuisance factor that requires control. Clearly, ideal queries and rational actions maximize system performance [2] and with respect to local performance, it is in the system's best interest to interpret noisy actions using available contextual information. In this view, interpretation is designed to overcome the searcher's shortcomings (e.g., by query expansion with synonyms) and to supplement the query with information the ideal searcher would provide given sufficient time and attention (e.g. the desired vicinity for the pharmacy)[1]. In effect, such methods tune the system to optimize performance on the information needs encountered most frequently, and as discussed above, the searcher is habituated to a simple set of actions. For needs that do not fit well with the tuning, the searcher knows few alternative means for clarifying, which leaves the model with insufficient information for effective interpretation. We speculate that where a system assists its searchers in process-learning, local search behavior will be more predictable across a broader range of needs, tightening the feedback coupling between the system and its user on difficult or rare needs. Research on this undertaking requires two major components: (1) developing a broader set of effective and useful alternative search actions and (2) developing effective methods for imparting process knowledge for these actions. We argue that the interdependent nature of interactive search requires evaluation paradigms that include the searcher's process-learning and that this approach leads naturally to deeper theories on user-centered performance evaluation and optimization. For example, conceptualizations of system failure may be advanced by more precise understanding of how searchers respond to failure and the factors that affect their responses.

1.5 Integrating learning into evaluation.

We close with a quick mention of how process-learning may be integrated into the overall evaluation paradigm. For system components designed to impart skill with a process, learning may be treated directly as an output or outcome for evaluation. For components involving novel displays or actions, prior learning may impede efficient use such that the utility of the solution is not gained without sufficient learning. In these cases, evaluation should include the learnability of the component. Where the focus is on more global processes such as searching-to-learn [10], evaluation may focus on more complex multi-factorial learning outcomes. More broadly, in any evaluation involving real human search interaction, process-learning is likely to be a factor. At a minimum, in experimental work where learning is a nuisance variable, learning effects must be controlled for in design and analysis.

2. REFERENCES

[1] Anderson, J.R. 1989. A theory of the origins of human knowledge. Artificial intelligence. 40, 1 (1989), 313–351.

[2] Baskaya, F. et al. 2013. Modeling behavioral factors in interactive information retrieval. Proceedings of the 22nd ACM international conference on Conference on information & knowledge management (2013), 2297–2302.

[3] Bateman, S. et al. 2012. The search dashboard: how reflection and comparison impact search behavior. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (2012), 1785–1794.

[4] Debowski, S. et al. 2001. Impact of guided exploration and enactive exploration on self-regulatory mechanisms and information acquisition through electronic search. Journal of Applied Psychology. 86, 6 (2001), 1129.

[5] Järvelin, K. et al. 2015. Task-Based Information Interaction Evaluation: The Viewpoint of Program Theory. ACM Transactions on Information Systems (TOIS). 33, 1 (2015), 3.

[6] Joachims, T. et al. 2005. Accurately interpreting clickthrough data as implicit feedback. Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval (2005), 154–161.

[7] Loraas, T. and Diaz, M.C. 2009. Learning new uses of technology: Situational goal orientation matters. International Journal of Human-Computer Studies. 67, 1 (2009), 50–61.

[8] Marchionini, G. 1997. Information seeking in electronic environments. Cambridge university press.

[9] McKay, D. and Buchanan, G. 2013. Boxing Clever: How Searchers Use and Adapt to a One-box Library Search. Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration (New York, NY, USA, 2013), 497–506.

[1] We ignore additional goals of modeling.

[10] Rieh, S.Y. et al. 2016. Towards searching as a learning process: A review of current perspectives and future directions Journal of Information Science. 42, 1, 19–34.

[11] Smith, C.L. 2015. Investigating the role of semantic priming in query expression: A framework and two experiments. Journal of the Association for Information Science and Technology. (2015).

[12] Smith, C.L. and Kantor, P.B. 2008. User adaptation: good results from poor systems. Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval (2008), 147–154.

[13] Turpin, A. and Scholer, F. 2006. User performance versus precision measures for simple search tasks. Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval (2006), 11–18.

[14] Vakkari, P. et al. 2003. Changes of search terms and tactics while writing a research proposal: A longitudinal case study. Information processing & management. 39, 3 (2003), 445–463.

[15] Vakkari, P. and Kekäläinen, J. 2011. An evaluation methodology reflecting the aims of search tools. Proceedings of the TPDL 2011 Workshop on User Oriented Evaluation of Digital Library Interfaces, Berlin (2011), 1–9.

[16] Wood, D. et al. 1976. The role of tutoring in problem solving. Journal of child psychology and psychiatry. 17, 2 (1976), 89–100.

[17] Xie, I. and Joo, S. 2010. Tales from the field: Search strategies applied in Web searching. Future Internet. 2, 3 (2010), 259–281.