Search evolution for ease and speed: A call to action for what’s been lost

Virginia M. Tucker*
School of Information, San José State University, San José, California, USA

Sylvia L. Edwards
School of Information Systems, Queensland University of Technology, Brisbane, Australia

*corresponding author: virginia.tucker@sjsu.edu, School of Information, One Washington Square, San José State University, San Jose CA 95192-0029 USA

Abstract
In recent years, leading website search engines have abandoned vital search features supporting complex information needs, evolving instead for the marketplace and for users seeking speedy answers to easy questions. The consequences are troubling, for researchers and for information science educators, with concerns ranging from the very relevance of search results and the unknowing of what is missing, to the novice searcher’s waning ability to frame potent queries and to learn ways to refine results. We report on a grounded theory study of search experiences of information professionals and graduate students (n=20) that contributes a holistic understanding of web searching, using its findings both to frame what is lacking in the design evolution of search engines for complex information needs and to outline a way forward.

One goal of the study was to evaluate an established model of web searching, called Net Lenses, a theoretical framework shown to be highly relevant during the study’s grounded theory secondary literature review. The original Net Lenses research used phenomenography to identify variation in the web search experiences of university students (n=41), evidencing four categories according to characteristics of searcher awareness, approach to learning, response to obstacles, and search outcomes. This study validated the model and led to an expanded version, Net Lenses 2.0, with five categories of search experience, reflecting the complex information needs of more advanced searchers. This resultant Net Lenses 2.0 model is discussed with its implications for search engine design, for advanced searchers and also for learning-to-search modes, much needed by searchers seeking to develop their abilities. The study’s implications coalesce in a call to action for more inclusive search interface design, and an agenda is put forth for how information researchers, educators, and literacy advocates can move forward in their intersecting domains.

Keywords
complex search, information behaviour, information needs, search engines, search interface design, threshold concepts
Introduction

As website search engines have evolved in recent years, they have moved away from vital search features supporting complex information needs, becoming more and more focused instead on being platforms for gathering data on users and on developing algorithms for speedy answers to easy questions. In a jarring illustration of what has been lost to users who need search tools for deep questions, the DIALOG search system was recognised last year with an IEEE Milestone Award, joining an elite group of inventions such as Marconi’s telegraph and ARPANET. The award citation notes that DIALOG was first to allow “iterative refinement of results” back in 1966, “preced[ing] major Internet search tools by more than two decades” (IEEE, 2019: sect. 12). It also preceded ARPANET, the forerunning technology to the Internet. Yet we find ourselves five decades later conducting most searches on systems that no longer support advanced iteration. Website search engines have no facility to refine initial search results, other than by limited options, such as file type, date range, and webpage segments (e.g., title, URL, text). Refining the initial result through inclusion of additional search terms is not supported, nor is faceted iteration or restriction of terms through proximity restrictions or field limiters, all standard fare since the 1970s on subscription database aggregator search systems.

Two critical search features previously available have actually been removed in recent years: word proximity operators and learning-to-search interfaces that supported the user in developing more advanced skills. That said, highly beneficial gains have been achieved in the intelligence and customisation of search, with search interfaces and algorithms now exceptionally good at supporting the middle majority of user information needs for quick answers. Marchionini (2019: 76) has stated that “today, searchers type or speak words, (the) phrases or questions are aided by autocomplete (anticipatory) assistance, and immediately are provided with sets of potentially pertinent items ordered in presumably useful ways. Alternatively, in faceted search systems, searchers select items from categories and navigate to desired items [consequently] the mental activity of search has shifted from problem statement (planned queries) to scanning and filtering results from occasional, conscious effort to frequent and automatic act(s) embedded in daily life”.

Underlying this design context is the motivating drive for the search engine developers: persuasive technologies. As described by Bates (2019), “The goal is no longer to make search results more valuable, the search industry is about making as much money as possible from people’s data,” and “people are not put at the centre of the information world these days” (‘How search really works’). Sponsored search results persist in prominent display at the top of results pages, and yet are less conspicuously identified as such. Further and much more insidious factors affecting the ranking of results are the well-documented racial, cultural, and gender biases in search algorithms (Noble, 2018), making the value of search results open to question, yet many users do not know they are present, much less to question them (Hansen and Rieh, 2016; Rieh et al., 2016).

Within the search industry, product design is focused more and more on the mobile experience, driving a more simplified search interface with fewer features, as well as a move towards more semantic search options (Oh and Tang, 2018). Another factor is gamification that exacerbates problems with how
content, particularly news, is presented (Noble, 2018: 154). Yet as we consider ways to go forward with
these search engine developments, what is the big picture of what we are gaining—and what have we
lost? True, if we can afford to invest in smart glasses, we may see our quick information needs even
more instantly met. On our walk from the office to the subway, our mini-screen will let us know the
trains are delayed and offer up alternative routes. In broad terms, however, the information
environment in which these advances and steps backward have taken place has been driven by the
primary objectives to meet the needs of the marketplace, not the human needs (Bates, 2019). The
consequences of what has been lost as search has evolved for the marketplace, ever-collecting the
minutiae of data about us, belongs in the conversations of information literacy advocates and in how we
think about future directions for search. The single search box is perceived by many as a deep dive—
some call it “doing research”—yet it can only perform in that way if the searcher knows tricks and
shortcuts to manipulate results. Even accessing the advanced search form has been made increasingly
difficult over the years (see Hard to Learn section below). The needle in the haystack cannot otherwise
be found, nor can there be confidence about what might have been missed. There is no meaningful
conversation (Winograd and Flores, 1986) with the retrieval engine and no longer the learning-to-search
mode to support a novice in discovering and practicing more advanced features.

What has been lost during the evolution of web search, leaving users behind, is discussed here and
illuminated through our study findings that explored the web search experiences of university students
and professional searchers, using a mix of methodologies to arrive at a holistic theoretical model for
web searching, Net Lenses 2.0. The research demonstrated a widening divide in approaches to search
and in understandings of how search works, representing a second form of ‘digital divide’. This divide is
between those who rely entirely on the search engine’s smarts, trusting in the relevance ranking, word
associations, personalisations, optimisations, and even the sponsoring of results—and those who can
tame the results through application of largely hidden ‘power’ features and strategies to overcome filter
bubbling and algorithmic biases (Fister and MacMillan, 2020; Hunt, 2020; Noble, 2018). In this paper, we
discuss the complex information needs and the praxes of advanced and professional searchers that have
less and less functional support from web search engines, using the study’s findings and resultant
theoretical model to present a holistic framing of these information needs. In addition, we use the
findings to show how search engines are losing ground in providing the learning-to-search modes that
are essential to furthering search literacy for complex information needs and critical thinking. We then
put forth solutions, involving both the restoration of features removed in recent years and new ideas to
move web search engines forward into designing for transparency and more truly interactive search
experiences.

With this landscape of search options gone missing as the problem statement for this study, three
research questions positioned the research design: What has been gained during the trajectory toward
simplified search? What has been lost to those users left behind? Where do we go from here?
Research context

With the critical importance of everyday search yielding relevant results well established, it is essential to understand the range of information needs to be met. Information literacy involves basic objectives so that everyday searchers learn how to assess search results, to choose authoritative over fake news sources, and to be mindful of privacy concerns and ways to protect search histories. A higher-level goal is to understand “research as inquiry” (ACRL, 2016), the very nature of engagement with information. This depth of engagement with information requires iteration in the interactions, both in order to understand and develop the question fully and to arrive at meaningful answers. Iteration in the form of query reformulation is only one approach to honing in on the actual information need. Questions and problems may range from helping everyday searchers to locate health information resources that are legitimate, to methods for protecting the browser history of domestic violence victims.

Yet even with the astounding accuracy of website search engines for quick answers, visionaries in information technologies are aware of what is lacking. Ramesh Srinivasan, plenary speaker at ASIST 2018, made a call to action for Google search results pages to have refinement options for users, including these three features: (1) an option to see how the results were ranked and ordered; (2) an option to see to what degree his previous search history affected the ranking; and (3) options for next steps in order to refine the results, beyond the current selections for time ranges and restricting to images, videos, or news (Srinivasan, 2018). In addition, without options to see and understand how the search results were generated, the learning that can take place during search is minimal. Likewise, without being informed both about how one arrived at the present step and about potential next steps, little learning is possible and little improvement in search skills and outcomes can happen. This is acceptable for quick answers, but not when deep and multi-faceted questions need to be answered. The research study reported on here provides detailed insights into the information needs of users who are performing searches with just these sorts of questions in mind.

The consequences of absent search options can be profound and costly, and they may be most glaringly evidenced in the reported experiences of researchers who are not experts in search, those who do not hail from the information professions or research domains. In a case illustration conducted by industrial designers that explored the search strategies of colleagues and practitioners in their own industry (Francis, 2004), researchers found that participants could not locate critical information relevant to their design projects, with the outcome being negative impacts on framing their project objectives and deliverables. Writing as a non-information professional, Francis locked onto the solution of controlled vocabulary as the key aspect that was missing in the search engines used in the study, recommending in her conclusions “a move towards developing Information Retrieval Systems founded on professional design jargon and terminology specific to the Industrial Design field” (2004: 650). Similarly, Brace (2012), in seeking solutions to tracking product requirements, put forth having vocabulary in the form of a checklist and concluded, “categorizing with the checklist is one method for supporting traceability and the application of a computer tool for an effective process” (2012: 77). Such findings from outside the LIS domain are particularly interesting for how they validate this study’s
model, which was based on the search experiences of students, from six of the eight academic faculties (departments) at the university (n=50), and those of LIS professionals (n=11).

A similar and prescient recommendation came from Winograd and Flores (1986) who proposed in their classic text that retrieval systems be developed to allow for recurrence in “conversations for action” (1986: 64), making it possible to interact as though conversing with “another person who shares our background” (p. 68). They provide support—again from outside the library and information science domain—for two components long known to be essential to search: firstly, interactive, iterative, conversational search, and secondly, vocabulary specific to the domain. Interestingly, at the time of their writing, the DIALOG search system had existed for over 20 years and offered all of these features.

**Market-driven, not user-centred, design**

In studying the workings of Google’s core search algorithms, Luh (2016) found that PageRank endures as the governing feature for ranking results; it places sponsored results first, with title weighting coming in second, and the URL and snippet following at roughly equal importance after that. The basis of ranking, in simple terms, was “the logic that the most linked-to webpages are most likely to be the most relevant...every link was seen as a digital vote” (Hunt, 2017: sect. 6). Needless to say, there has been significant work—and ongoing work—to improve search algorithms and to introduce enhanced semantics, not only into search engines, but also into semantically tagging documents (Siemens AG, 2017), all in attempts to more correctly interpret common language queries, to offset the weight of the ‘digital vote’, and ultimately to improve the odds of providing the user with the answer they are seeking. However, these changes have come simultaneously with far-reaching and market-driven concessions on privacy and user control over results. Privacy agreements may be viewable with effort but are minimally modifiable—with even greater effort. User control options have lost ground to search algorithms that give priority to the advertising revenues that drive search engine design.

The broad view may be summarised as that the search engine industry has shifted radically from being centred on users’ information needs to being driven by market forces (Bates, 2019); indeed, the core business objective is turning a profit from people’s data, whether it is their search histories, location tracking, or purchasing patterns (Bates, 2019). Users and their personal data are what Google is selling; we have known for some time that it is not selling a search product (Vaidhyanathan, 2012). Both speed and mobility are key driving forces, fuelling human compulsions to search more and click more, which then fuels the generation of more results, more advertising links displayed, and more revenue. Robertson (2016) has characterised these developments as feedback loops that have shaped the search experience. In addition to the more obvious loop connecting products and users, Robertson described the tightening loop around designers and advertisers as “each [is] interacting with (influencing and being influenced by) the search engines” (2016: para. 2). Carrying this broad view of market forces further, conflicts of interest are evident in search ecosystems, most visible in explicit labelling of sponsored search results, yet also hidden in artificial boosting of results ranking through metadata manipulations and the lack of governance for algorithmic decision making (Just and Latzer, 2017; Latzer et al., 2016).
This reality has wide ranging repercussions. In his call for improvements to search tools for research purposes, Crotty (2019) wrote about issues of access in developing countries and also about gaps in discovery capabilities in the search options that restrict essential access to research literature that is critical to areas such as medical information. Crotty reflected on the Ebola crisis in 2014 and how blame for gaps in research content was placed on subscription database paywalls, when in fact the needed research literature was available on open access platforms. In the oversimplification of open vs. closed access, problems with discovery capabilities and search tools were swept under the proverbial rug. As described by Schonfeld (2015), the starting point for many researchers is not their library’s database services, but the “internet services that enable our use of multiple devices anywhere and effective switching between them” (2015: 2) and their enabling and anticipatory discovery features. Noble (2018) has warned, “As information moves from the public sphere to private control by corporations, a critical juncture in the quality of information available and the public’s ability to sift and use it is at stake” (2018: 153).

**Anticipatory search and algorithmic bias**

In short, the strident prioritisation within the search industry of revenue over relevance is evident in the interface, the search options available, the search modes, and the ranking of results. Critically, it is also evident in the algorithmic biases, well documented and present in so many forms: racial, cultural, and gender, to name a few (Noble, 2018; Tufekci, 2015). For the context of this study, algorithms are defined as “computational processes that are used to make decisions of such complexity that inputs and outputs are neither transparent nor obvious to the casual human observer,” a combination that “raises questions of accountability, process, and transparency” (Tufekci, 2015: 206). Anticipatory search perhaps demonstrates most glaringly the effects of algorithmic bias, captured by the search string, ‘why are black women so...’ and its predictive completions, on the cover of Noble’s 2018 book, *Algorithms of Oppression*. Algorithms at their best support shortcuts to search results, anticipating most likely best answers based on the searcher’s own location and previous searches. At their worst, they predict answers based on hidden factors, such as advertising sponsor influences.

Very recent changes have made differentiating sponsored from unsponsored results even more difficult (Ng, 2020; Shieber, 2020). Significant impacts go far beyond Google search results and Facebook feeds, however, and “algorithmic gatekeeping” (Tufekci, 2015: 206) is present in almost every large digital dataset, from human resources records to government files. The complexities of algorithms are outside the scope of this paper, yet a simplistic perspective is included here to represent both the importance of their impact on search results and, more specifically pertinent to this study’s context, the fact that most users are unaware that they exist or to what extent their results are affected. Building into the learning-to-search experience an awareness of algorithmic effects, a way to show/hide the most impactful factors that are influencing the search results, would be a step towards transparency. Braun (as cited in Fister and MacMillan, 2020: 37) sums up these concerns,

> These algorithms pervade nearly every domain in our lives, and in that ubiquity, we are subconsciously challenged to flex many different literacies—information, data, visual,
statistical, and technological literacies, to name a few—to interrogate the intentions of those systems, how they are constructed, and how they control our access to information. Whether we realize it or not, this continuous assault can quickly fatigue our ability to question those systems themselves, tightening the control they have over us even further. In the face of this engineered reality, it can feel easy to become nihilistic about the agency we exercise in responding to these systems as they reconstruct our daily lived experience. Perhaps, the most radical act we can perform, then, is to reject that nihilism and recognize our role as actors in exposing how these systems deprivilege complex, nuanced, and situated knowledges.

Providing a broad view of algorithms in the wild, Zittrain (2019) has written about the effects of machine learning systems that are dependent on automated algorithms, telling how they have moved from their early research domains in medicine, for example, and into applications involving information retrieval, online retailing, and the stock market. Zittrain refers to the “intellectual debt” that is accruing from this process of our reliance on algorithm-generated results, cautioning, “As we begin to integrate their insights into our lives, we will, collectively, begin to rack up more and more intellectual debt” (2019: para. 4).

**Hard to learn—and getting harder**

Uncannily, as users’ information needs continue to become more complex and sophisticated, search features are being removed. In addition, interactive modes that supported the learning of search techniques have also been taken away. As an example, and alongside the loss of proximity operators (other than exact phrase) mentioned earlier, Google removed its learn-to-search mode that displayed the syntax for advanced options on-the-fly, creating a bridge from basic search into options such as webpage section delimiters and file type and site name restrictions (Figure 1). This made it possible for novice searchers to experience and to learn—in the moment—more advanced techniques for searching and ways to refine and/or expand their results, thus developing the novice’s abilities and, at the same time, enhancing the search results.
Figure 1. Google Advanced Learn-to-Search Mode, circa 2008. (source: Blachman, 2015)

Figure 2 illustrates using the search form with required terms, OR Boolean logic, and excluding terms (NOT logic). The single-line format of the search was automatically generated by the system in top search box, helping the user learn about the logical operators.

Figure 2. Learn-to-Search Mode Illustration with Boolean (logical) operators (source: Blachman, 2015)

The learn-to-search mode (Figure 1) further provided an in-context and seamless visualisation for the searcher that displayed shortcuts and power searching options directly in the basic search box, connecting the searcher with knowledge about next steps and suggestions relevant to their information need. Hunt (2017; 2020) has reached beyond these logistical enhancements, putting forth a search engine prototype that actively engages in conversation with the searcher following a Socratic dialogue form, with the objective of “consciously enacting modes of thinking upon internet information retrieval”
Studies show that university students have a growing appreciation of search strategies as well (Martzoukou, 2008).

Google Scholar, designed for researchers and higher education students, has its default interface, as well as an advanced search form with basic Boolean options, searching by all the words, exact phrase, at least one of the words, or to exclude a word; and the ability to search within the article vs the title, to search by author, and to apply date limitations. Interestingly, Google’s advanced search features include more options than Google Scholar, such as restriction by language, region, last time updated, file types, usage rights, and site or domain limits. Its ranking parameters are most heavily weighted toward how often and how recently an article has been cited and where it was published (Google Scholar, n.d., “About”; “Inclusion Guidelines”), useful to academics, and yet it is lacking in advanced and iterative search options (Halevi et al., 2017). In recent research by Alotaibi and Johnson (2020), a key conclusion was that their study “enable[d] our understanding to go beyond the assumption that ease of use and convenience are the main drivers behind the decision to use [Google Scholar]” (2020: 14).

Google has also recently added a search platform for datasets (Google Webmasters Support, 2019), a promising tool for isolating specific datasets that are publicly available for research. This joins Google Trends and Insights, highly useful tools for analysing social media and web query patterns. These advances cannot be minimised, and yet they have not come without costs—some hidden, and some more conspicuous. Much of the search functionality is sufficient for webpage searching, but the major search engines added content that requires more extensive search capabilities. Google, with the lion’s share of the search engine market at between 63% and 92%, depending on country and the metrics used (Price, 2020; StatCounter, 2020), leads all other search engines. Yet the Google customised searching experience comes at a price for its users; most notably, in advertising inundation and privacy intrusions. Less conspicuous are factors unknown to most users, primarily the filtering of results based on parameters such as sponsoring levels, previous searches, location, and algorithmic bias (Bozdag and van den Hoven, 2015; Noble, 2018; Pariser, 2011; Vaidhyanathan, 2012, 2018). Searchers are often unaware of differences in curation practices between subscription-supported content and open internet content. Users make assumptions about what is being searched, and yet the differences between the highly curated databases from subscription vendors and open web content are as vast as those between organised and indexed materials within an ages-old library and the heterogeneous landscapes of the Wild West.

Fifteen years ago, Nielsen (2004), in describing the phenomenon of “information snacking” (‘Answer-focused search’), asserted that:

A major change over the years has been a declining emphasis on using search to identify good sites. Rather than hunt for sites to explore and use in depth, users now hunt for specific answers. The Web as a whole has thus become one agglomerated resource for people who use search engines to dredge up specific pages related to specific needs,
without caring which sites supply the pages. Search engines have essentially become answer engines. (Nielsen, 2004: ‘Answer-focused search’)

A decade later, Laubheimer (2016), with the Nielsen Group, stated:

One of the great insights of the original Google team was that precision is more important than recall for most web searches: there’s so much information available on the Internet that nobody would even want to see all the relevant results. For web-wide search, it’s much more important to focus on high precision: to make sure that all the top 10 hits are as highly relevant as possible to the user’s current problem. (Laubheimer, 2016: ‘Two metrics’)

With this development focus on finding precise answers to specific questions, functionalities in support of in-depth questions have been side-lined. In addition, and integral in this trajectory, there is a more fundamental and worrisome concern: that novice searchers no longer develop the ability to think critically before entering terms into the basic search box or indeed to pose well-framed questions. Because there are assumptions in the search engine’s algorithms about logic and proximity applied to the terms (among other operations), the searcher bypasses the contemplation involved in framing a potent query. “Google is known as a search engine, yet there is barely any searching involved anymore. The gap between a question crystallizing in your mind and an answer appearing at the top of your screen is shrinking all the time. As a consequence, our ability to ask questions is atrophying” (Leslie, 2015: para. 4). The abilities to formulate an in-depth question and to plan for search terms to be entered—whether for a middle school paper or a dissertation—go hand in hand in support of critical thinking skills and learning to consider information authority, credibility, and context. Russell, a research scientist for User Happiness at Google, recently wrote, “Online information often doesn’t come with much context. You have to supply your own or know how to find it. Just as important, you have to understand—deeply understand—that context is just as critical as the information you seek” (2019: 295).

In addition to the practical effects of these missing interface elements that support learning to search, these gaps reinforce corresponding gaps in user mental models (Han et al., 2020) that have been documented and which have a profound impact on user ability to scale the learning curves connecting novice search skills to higher proficiency in search for more complex information needs. The types of errors made by novice searchers (Ondrusek et al., 2019) can be overcome by in-context assistance from the search engine, whether in the form of context-sensitive help or learning modes (see Figure 2), and discussed in the literature (Eickhoff et al., 2015; Fessl et al., 2018; Moraveji et al., 2011). During the activities of searching, a user is intensely receptive to learning new ways, both to better understand the processes of search as they are occurring and to refine and iterate in order to improve the search outcome (Ghosh et al., 2017, 2018; Theng et al., 2016). Yet users are instead learning to become “ever more reliant upon the technologies’ powers of reasoning and less reliant upon our own” (Hunt, 2017: sect. 9).

The study’s findings provide a model for understanding the highly proficient searchers whose needs have gradually been losing ground in the design of search engines as they have turned toward the
majority and toward becoming answer engines. Additionally, the loss of learning modes to support furthering the search skills in the moment of search activities can be better understood through the research study’s structured exploration of how students engage with information. This includes, for example, how they approach learning to search, what they focus on, and their responses to obstacles during a search. Ultimately, search interfaces need not be either-or, but rather they can be designed to meet the information needs of both novices and experts (Vega, 2017). We advocate for this combined search engine design for these very reasons. The information environment is in fact ripe for a win-win: Restoring an alternative and intelligent search mode that supports user learning about search and real-time development of search skills would advance, not only the user’s ability to manage complex information needs, but also the search industry’s need to reach searchers whose needs are no longer being met in the recent evolution of search design toward quick answers.

Research design and methodologies

The original Net Lenses study (Edwards, 2006) used phenomenographic methods (Marton, 1994, 1997) to uncover variation in students’ experiences of web-based information searching. Incorporating first year, third year, and postgraduate student perspectives, the participants (n=41) came from six of the eight faculties at QUT (Queensland University of Technology). Different cultures, ages, and genders were represented. The students were asked during interviews to describe a recent search experience, and to describe how they learned to search for information using various web-based tools. Approximately one-third of the participants were also videoed as they undertook an information search using a think-aloud protocol. Analysis involved an iterative process of seeking meaning and structure. Four categories of explicit variation were discovered, with each category described using words of the participants. Each category is characterised in terms of referential and structural components constituted in terms of the critical dimensions of variation, including focal elements, approaches to learning, and reflective practice. The resulting model, called Net Lenses, has been presented previously (Edwards, 2006), and validated in multiple subsequent studies (Edwards et al., 2002; Edwards, 2004; Edwards, 2006; Edwards et al., 2010).

This study, aimed at further evaluating the Net Lenses model, used constructivist grounded theory methodology (Charmaz, 2014; Davis, 2014; Dey, 1999; Tucker et al., 2016) to explore the search experiences of 9 highly proficient graduate students in an MLIS degree programme and 11 experienced search professionals from around the United States and Canada (total participants=20). The datasets (Tucker, 2016, 2020) were captured search tasks with think-aloud protocols, and semi-structured interviews, conducted before and after the searches (Branch, 2000; Charmaz, 2003). (Refer to Appendix B: Study participants, search tasks, interview questions.) As a grounded theory study, Tucker’s research examined relevant extant theory in the secondary literature review and later stages of the work. The threshold concepts theoretical framework was the primary extant theory that informed the research; the secondary extant theory was the Net Lenses model. In order to rigorously validate and contribute to the Net Lenses model, the datasets (captured search tasks and semi-structured interviews) were re-examined in a new phase of grounded theory thematic analysis. These results, reported below,
ultimately validated and expanded the Net Lenses model, making it more inclusive of highly advanced searcher experiences.

Results
To further demonstrate what is missing for the users left behind in the recent evolution of search, we report our research on how web searching is experienced, evaluated according to four parameters of the user’s engagement with the search engine: (1) awareness structures (such as search tools, information quality, and the user’s worldview), (2) approaches to learning while searching, (3) assumptions about and responses to obstacles during search, and (4) search outcomes. These parameters were at the centre of the Net Lenses model developed by Edwards, that established four categories for how web searching is experienced by university students, with the parameters described above defining the boundaries of each category. Each category was described in vivo, using words of the study participants for how they experienced information searching:

Category 1. As looking for a needle in a haystack.
Category 2. As finding a way through a maze.
Category 3. As using the tools as a filter.
Category 4. As panning for gold.

For example, in Category 1, searching was described as looking for a needle in a haystack, highlighting the lack of awareness in the structure of the environment being searched or the quality of the content. The participants in this experience category were inherently frustrated by the searching experience; whereas in the remaining three categories, searchers demonstrated an awareness that the environment is structured to assist them while searching. Within each category, meanings were assigned to the search experience, and within each of these experiences, there were four parameters that characterised the experience: awareness structures, approaches to learning, responses to obstacles encountered, and search outcomes. For example, the two higher level search experiences, Categories 3 and 4, included an awareness of the quality of information and, in their approaches to learning, users were likely to include some planning around and analysis of the information need. Category 4, the most advanced group, was triggered by the discovery of differentiations in searcher awareness of the “quality of information” and a “quest for improvement”. While the study identified levels of sophistication in searching behaviour, with the more discerning search strategies evident in the higher levels, it also demonstrated a variety of lenses through which university students may view, experience, and learn about web searching. They can choose the appropriate lens for a given search task, and swap between lenses as needed. A detailed summary of the four categories is provided in Appendix A.

In this new study, in order to fully explore the expectations a user has of the search engine when working with complex information needs, these same four parameters were enlisted to study the experiences of searchers transitioning along the continuum (Dreyfus, 2004) from novice and proficient searching towards expertlike searching. The study analysed search praxes and conceptual knowledge, as well as traits of the searcher when encountering obstacles, new learning, and different outcomes. The
findings from the study ultimately resulted in a fifth category for the model, which we now call the Net Lenses 2.0 Model. Notably, Edwards had speculated in the original study about categories that might exist “above or below those already identified” (2006: 192) if using a different set of participants, and that this was an area for further research.

This study’s datasets that had originated out of the PhD research of Tucker were re-analysed in order to explore how they informed the extant relevant theory of Net Lenses. In the information disciplines, established models of information behaviours have often been extended by other researchers, leading to enrichment of the original works and further validation of the model. The berrypicking model of Bates (1989) is such an example. Lillard and Ha (2015) highlighted the research emanating from the original berrypicking work, examining the milestone years 1989, 2002, and 2005, and citing the early search expertise work from Tucker’s research (Lillard and Ha, 2015: 115). In a similar fashion, Hsieh-Yee (1993, 2001) and others built upon and significantly extended the work of Bates (1979, 1987) on search tactics and their applications.

For the datasets, the focus was on professional and highly adept searchers, which strongly positioned them for the higher levels as the research frame, so that further development of the Net Lenses model was highly relevant as a secondary theoretical analysis stage. Half of the 20 participants were experienced information professionals whose positions demanded advanced search skills, and the other half were graduate students who had been selected by their instructors in an advanced search methods course for having demonstrated expertlike search abilities. (See details in Appendix B.) This study population was well situated for studying the outlier edges of the Net Lenses Category 4.

With the further context that Tucker’s research outcome had been a grounded theory for search expertise that encompassed critical (threshold) concepts, traits, and praxes, our new study was positioned particularly well. For this study, Tucker re-analysed the data in two stages, then evaluated the results in terms of how they could enrich the established Net Lenses model. The first stage of analysis involved the coding categories that were relevant to the model. The resultant themes and primary coding categories from this first stage are summarised in Table 1. The summary addresses the attributes of the expertlike search experiences and what they required of the user: conceptual knowledge, as well as traits and praxes (such as tools applied, skills, and approaches or styles of searching). The components of conceptual knowledge are restricted to threshold concepts (Meyer and Land, 2003), concepts that involve “learning to see some aspect of the world in a totally new, transformative, and often counter-intuitive manner” (Tucker et al., 2014: 150). It is important to note that praxes and traits do not have the characteristics of threshold concepts, however, the praxis attributes were critical to understanding the nature of expertise among searchers, and they were key factors in indicating the boundaries leading to Category 5 in the Net Lenses 2.0 model. For example, the threshold concept of ‘information vocabularies’ includes concepts such as natural language and controlled vocabulary, as well as praxes for fluidly using language-based tools like word proximity restrictions and subject domain terminology. The most profound attributes of expertlike searching were
in the threshold concepts involved with the ability to fusion key concepts and to integrate praxes in ways that characterise expert performance (Berliner, 1994; Schön, 1983, 1987; Simon and Chase, 1973).

Table 1. Stage 1 Analysis of themes and codes: Concepts, praxes, and traits relevant to Net Lenses.

| Theme                        | Coding categories (Concepts, Praxes, Traits)                                                                 |
|------------------------------|-------------------------------------------------------------------------------------------------------------|
| information environment and domain knowledge | concern re. misunderstanding info environment C |
|                              | conceives reference interview from broad view C                                                               |
|                              | perceives total environment, including sources C                                                                |
|                              | deep knowledge of provider practices P                                                                         |
|                              | has rapport with search engine P                                                                               |
|                              | subject domain knowledge P                                                                                    |
|                              | if applicable, has rapport with client/organisation P                                                          |
| information structures       | perceives info structures from broad view C                                                                    |
|                              | grasps transparency of info retrieval system C                                                                   |
|                              | visualises structure of search C                                                                               |
|                              | conceives cited reference connections P                                                                         |
|                              | understands database and record structures P                                                                     |
|                              | grasps fields as important, weighted C                                                                          |
|                              | understands and applies term weighting and frequency C                                                          |
|                              | unpacks the topic, conceives building blocks C                                                                   |
| Information vocabularies     | natural language, keywords C                                                                                  |
|                              | controlled vocabulary fluency C                                                                                |
|                              | proximity relationships P                                                                                     |
|                              | uses thesaurus adeptly P                                                                                       |
|                              | subject domain terminology knowledge P                                                                        |
|                              | word-term fluency, truncation, synonyms P                                                                       |
| fusion of concepts and integration of praxes | integrates search tools C, P                                                                                   |
|                              | synthesises information C, P                                                                                   |
|                              | integrates results in the moment C                                                                             |
|                              | knows and applies different styles of search C                                                                  |
|                              | learns new systems easily C, P                                                                                  |
|                              | adventurous in trying different approaches T                                                                    |
|                              | perseveres when stymied T                                                                                      |
|                              | enjoys the hunt in the extreme T                                                                               |
|                              | light on feet during search, flexible T                                                                         |
|                              | reflects on search outcome and process T                                                                        |
|                              | identifies with being a searcher T                                                                             |
|                              | anticipates and visions the stages of search T                                                                  |
|                              | knows when to stop a search, call it done T                                                                    |

14
The second stage of working with the data was analysing the codes and themes according to the model’s four parameters: awareness structures, approach to learning, response to obstacles, and the characterisation of search outcome. The results, showing the analysis for how the data informed the Net Lenses model, are summarised in Table 2. Here, the highest level of web search experience in the extant Net Lenses Model, Category 4, is differentiated across the four parameters from the emergent Category 5.

*Table 2. Differentiating Net Lenses Model Category 4 ‘Panning for gold’ and new Category 5 ‘Dancing’.*

| Parameters of Search Experience | Category 4 in Net Lenses Model | Attributes Suggestive of a Category 5 Search Experience | Stage 2 Analysis of Search Expertise Datasets: Coding and Themes |
|--------------------------------|--------------------------------|----------------------------------------------------------|---------------------------------------------------------------|
| Category’s characterisation    | “panning for gold”             | “dancing in the search environment”                       |                                                               |
| Parameter 1: Awareness structures (searcher’s focus) |                                 |                                                          |                                                               |
| information environment        | strong awareness               | profound awareness, extending to understanding of content provider practices | threshold concept: information environment                     |
| tool structures                | strong awareness and use of structure of tool | understands how tool works (e.g., relevance ranking algorithms); may customise tools | tool structures within threshold concept: information structures; praxis: combining different tools |
| information quality            | strong awareness; interested in primary, rather than secondary, sources | awareness extends to connections and nuanced differences among sources; knows how content creators have affected information quality | information quality within threshold concept: information environment |
| Parameter 2: Approach to Learning |                                 |                                                          |                                                               |
| confidence with IT generally   | high; aware of possible mistakes, more likely to self-correct, ask for help | high; may troubleshoot to determine cause of problems with search | coding categories: analysing, evaluating, visioning, anticipating |
| use of search planning         | planning is evident; may be written down before searching; often analyses term and synonyms before searching | planning is extensive; may look up terms pre-search, use thesauri and codes; consult with other professionals; collaborate and share strategies | planning is routine; terms and contingencies often jotted down pre-search; analyses terms, synonyms, database vocabulary information vocabularies, threshold concept coding categories: planning ahead, unpacking topic, visioning, anticipating |
The new Category 5 was labelled “dancing in the search environment,” a composite term based on the participants’ descriptions of the extreme levels of flexibility required when searching for complex topics and when using multiple types of sources. For example,

*Being able to modify in real time is important.* [P-14]

*Being able to handle moving targets, changing directions.* [P-17]

They were highly adept at searching, but their engagement went far beyond the activity of search; they were keenly aware of information in the moment and how their strategies might affect outcome, reflecting on the processes of iterating and refining and learning (in addition to frequently contributing content themselves). These experiences and approaches were evident both during the screen-captured searches and in the interviews conducted pre- and post-search.

Category 5 was differentiated from the Category 4, “panning for gold,” by the participants’ ways of describing their search experiences, with particularly critical differences evident in how they experienced engagement with the information environment, their roles in it, and their awareness of processes occurring before, during, and after search activities. These processes typically included a deep awareness of content creation and maintenance that was not present in Category 4:

*Key concepts were really understanding the content you were searching. This is something that's absent today, people have no idea what they're really searching. Second, structure of databases, fields, then the various connectors and tools the system has. It's really the marriage of those three things. Content, structure, tools to query.* [P-13]
Their reflective activities relevant to learning were a further differentiation, and these reflections took place both in real-time during a search and after searching.

*I think that here in the Google age where people often make perceptions so they make up for deficiencies. But you always need to think about how something might be written about in an article. That can be a hard concept. The way search engines are designed now, sometimes the engine complicates that.* [P-14]

In many instances, participants discussed the roles they held in careers, such as developing database products, educating others, and/or being an intermediary for others with complex information needs, in formal or informal settings. One participant described their role as intermediary in this way:

*A searcher needs to understand how to ask questions and how to help their [client] articulate what an acceptable response looks like, trying to understand the landscape.* [P-16]

Another described their role on a product design team and the insights they were able to contribute:

*Having been a user of their systems, it made me, I hope, think of things more from a user perspective, than a lot of the people that were designing the systems.* [P-14]

The characteristics of the new Category 5, outlined in Table 2, may be explicated further according to the Model’s four parameters:

1. **Awareness structures:** the searcher’s focus during the web search experience. The ability to discern what was most critical to success in the search was evident only in the Category 5 searchers. For example, there were choices made about the sources selected (Chu and Law, 2007; Walton et al., 2019) and about how rapport with the search engine affected outcome (Oh and Colón-Aguirre, 2019; Okhovati et al., 2017).

   *There was a comfort level to be ascertained with the system itself. And I think that's true in any search system.* [P-09]

   *You need to know what's been devised to help you separate the kinds of things you want. All of that is part of expert searching, knowing the content and knowing the interface and its tricks.* [P-15]

2. **Approach to learning:** how the searcher approaches learning about search protocols and options before and during the search, as well as how they reflect on the learning afterwards. A high level of willingness to experiment while learning about a new system or subject domain content was evident in the participants (Duarte et al., 2015; Fessl et al., 2019). They were highly experimental and open to exploring, while also being analytical and reflective of the learning processes:

   *I like to think of it as a mystery, and I have to get to the bottom of it.* [P-06]

   *I'd go back and look. What is causing this? Is it the terms I'm using? Am I using a limiter that doesn't exist in another file? Am I doing something wrong? Should I be changing it? And, how much does it really matter?* [P-12]

They also spoke of collaborative approaches to learning, whether with colleagues or other professionals and, importantly, of enjoying the process immensely.
The fact that I find it fun motivates me and has definitely taught me a lot because, anytime you’re doing your search, especially if it’s for a type of information or resource you haven’t searched for before, you learn from the process. [P-12]

3. Response to obstacles: the degree to which the searcher persists when obstacles or dead-ends are encountered during a search, and how willing they are to explore new avenues and adventure into unknown possibilities for resolving the obstacle.

   It was knowing when to stop, but also knowing how many options to exhaust before you stop, because you should exhaust quite a few. [P-06]

   It’s more of a process. You can find a good stopping point and then continue exploring, which may take you in another search direction. [P-08]

4. Search outcomes: Participants exceeded the Category 4 characterisation of “panning for gold” by becoming expert enough at searching to have a highly tuned sense of “knowing where the gold might be buried” [P-19]. This was not only the result of years of experience with different resources but, for the novices in the study participants, an ability to see the patterns in information structures and to discern the likelihood that a resource would be worth exploring for a complex topic. One participant took this further into how they supported others in being able to develop their questions:

   I’m expected to know the places to go to find the answer or even to anticipate how they might want to frame their question. [P-16]

Finally, because the study looked at the nature of expertise among searchers, the outlier aspects of the Net Lenses model, described as the internal and external “horizons” (Edwards, 2006) were considered as well. For example, Parameter 1 (structure of awareness) has internal and external horizons—it is as if the searcher’s view (or lens) or awareness consists of some parts seen clearly and in focus (the internal horizon) and other parts that are less clear (the external horizon). These horizons help to define the boundaries of the different categories. It was further important to narrow the focus on the elements existing in the Category 4’s external horizons that were of particular interest to the characteristics of search expertise that emerged. In addition, the experience of reflecting on the strategy of searching was important, as was refinement of the strategy. This reinforced the finding of reflection as a common practice among experts (Schön, 1983, 1987).

Findings and discussion
In examining the findings from this study, we first consider its insights into web search experiences and next how these insights can inform search interface design decisions going forward. In linear programming we know mathematically that we can have local maximums and global maximums. In less mathematical terms, it is the idea that when we have climbed the mountain and seen the top of it, it is only once there that we see the next mountain beyond, the one that is higher than the peak we have just climbed. The first peak is one of many local maximums, but the peak beyond maybe the final global maximum, e. g., the Denali of the Alaska Range. To relate that to these research findings, we often
experience that, while we have searched for something—an experience made easy for us by today’s search engines—what might we have missed in the local maximum we have just reached, thinking it is the global maximum? What might be beyond that we are unable to see?

In the original Net Lenses study, some participants were able to use all four lenses and select the appropriate lens to begin a web-based search, but in other cases they might only have two or three of the lenses through which they could operate. Importantly, it was clear that the student’s prior experiences influenced how they searched. This aspect was the most concerning, because if users only participate in an overly simplified search experience, they are unlikely to learn how to search for higher quality resources and in more advanced ways. It was also postulated that, when something is truly difficult or almost impossible to find—in other words, when it is truly a needle in a haystack—then the only the advanced search skills of a Category 4 lens will enable the student to find that needle. These skills characterised Category 4: the ability to refine and reflect on results, to improve the search strategy, the ability to understand and seek quality results, the ability to ‘pan for gold’; to grasp that only this combination of abilities will find that genuine needle in a haystack or, more accurately, all the needles hidden in the hay. The problem is, as search engine interfaces become increasing simple, being designed for novice searchers, we are removing what is made possible by advanced search interfaces and output options. In overly simplifying the interfaces and removing options for refining search results, we have also narrowed the user’s field of vision to what is within easy sight and likewise lost an aptitude to wonder what might be missing. The searcher can see only the immediate surrounding territory and not the information landscape that lies beyond, nor do they have a sense that it is there.

The argument here is not to abandon simple search interfaces. They are powerful tools that, by and large, provide the simple and quick answers often needed yet, as described above, they can constrain what a novice is able to learn about advanced search methods and experiences. What we have lost is what researchers need and may take for granted: locating quality research and the assurance that we have been thorough in our search. In simple web search interfaces, we may be reaching only the local maximums in the information environment. The majority of users may be content with that result, as they have found bronze and silver—but have they found the primary resources, not the secondary resources—have they found the gold? We must reinstate easy access to advanced searching features, as well as access to the modes supporting ways to learn about these features. Web search experiences should display the options to reach for quality resources and the functionality to iterate in refining the results, if we want people to learn to look for the global maximum. Enhancing the transparency and human-in-the-loop elements of search engines, along the lines of ‘explainable AI’, would enable users to learn to be better at searching, at formulating queries, and understanding ways to overcome the system’s built-in biases.

Two decades ago, usability guru Jakob Nielsen criticized the search interface design developments that were focused “purely at casual users. In-depth content and advanced information should be added to sites to provide the depth expected by experts” (2000: 1). In this early warning, Nielsen signalled that more and more users were becoming expertlike in their search behaviours, adding that we “must study
users over time as they develop expertise in using the site or service” (p. 1). Writing ten years later, he pronounced with apparent relief that “the balance between novice and expert users tilts in the direction of the experts” (2010: 1). Now we find ourselves another ten years later with the design mindset having tilted back in the direction of novice searchers, with search engines focused on supporting speedy answers above all and, second, anticipatory answers.

The intersecting studies reported here—using phenomenography and grounded theory, and a wide range of participants from first-year university students through to highly proficient post-graduate students and practicing search professionals—are rich sources for an understanding of new directions for search engine design that are both rigorous and holistic. The research findings, and the new Net Lenses 2.0 model that was the outcome, provide a renewed call to action for refocusing search interfaces to be inclusive of expertlike searching and complex information needs. Through these findings, the Net Lenses model was both further validated and enriched by Tucker’s research through the establishment of a fifth category for the model. The new Category 5 bridges the gap between highly capable graduate students conducting searches and information professionals; it illuminates both their information experiences and what is needed from search engines to support their complex information needs. The advanced searchers participating in the study demonstrated praxes, traits, and understandings of information concepts that had characteristics of expertise and which were not supported by the web search engines of today.

Additional implications of this research include the learning that is possible during search activities, when and if the search tool is designed to support it. In Marchionini’s words (2019), “Humans are learning systems” and the “skills and strategies (we use to search are developed rather than acquired” (2019: 81). More importantly, he contends that search is a “subtask of learning” (2019: 82), and we learn how to search, and make sense of the searching process, while we are searching. Marchionini calls for the IR research community to “build bridges to the neuroscience and behavioural science research communities so that we can understand the biological, affective and behavioural consequences of ubiquitous search on the whole individual and society at large” (2019: 84).

The search features most needed to support complex information needs are ones allowing the user to iterate on the results, to expand and to narrow. Faceted search options (restricting by document type or publication date, for example) address some of these needs to narrow results, however, further transparency about options to selectively switch-off personalisation (Salehi et al., 2019), for example, is also called for, as is re-ranking (sorting) of results by user-specified parameters. Second, an intelligent search mode that supports in-the-moment user learning about search and development of search skills would move users from novice skillsets to more abilities needed to manage complex information needs. Such enhancements could also promote the search industry’s own objectives by making it possible to conduct searches for more than quick answers.
Conclusion

In reflecting on this year’s IEEE Milestone recognition of the DIALOG search system and how it pre-dated the earliest Internet search engines by more than two decades, we also recall the defining characteristic of a DIALOG search experience that gave it its name. Summit (2002) tells the story of how, on a long road trip from the San Francisco Bay Area to Oregon, his wife Ginger remarked on how the breakthrough of his team’s invention was that it was like a conversation with the database content. Like having a dialogue, in other words, she said—and thus the name was chosen. This capability to conduct a conversation during search, to iterate and refine, to explore results and either extend or restrict them, based on what was displayed, was the essence of the first interactive search engine in the 1960s. This capability has endured as critical to conducting searches for complex information needs, as evidenced in the research reported here. In addition, the ability to overcome the filtering and algorithmic biases of web search engines is essential if searching the rich content—now accessible along with the non-authoritative content—is to be reached, if the user is to have a chance at perceiving what’s beyond the local maximum, that initial page of results, and, more than this, beyond the basic workings of the search engine as a tool. These improvements—along with restoring features removed in recent years, such as learning-to-search mode and more nuanced proximity connectors—are essential to meeting the information needs of researchers and also those of everyday searchers whose questions may go beyond the easy and quick.

Ultimately, we have gained free access and easy access, and yet critical elements of refining and iterative search options have been lost or so buried that most users are unaware of them. Simultaneous with this evolution, learning-to-search modes and in-context help for advanced search options have also fallen away, leaving searchers looking to grapple with complex information needs without ways to learn about what might be possible. As market forces drive search engines to support searching for easy and speedy answers, the opportunity to serve users with complex questions—and they are the future—is not to be underestimated. Fully grasping this win-win potential, both for users and for the search industry, is a rallying cry.

Acknowledgments

The authors gratefully acknowledge the participants in the research, their candour in speaking about their search experiences, and willingness to be observed during search activities. Critical friend feedback from Drs Laura Anderson, Debra Hansen, and Roger Summit supported refinements to our manuscript and was immensely appreciated.

Declaration of Conflicting Interests

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. Ethics clearance and permission to use participants’ anonymised
datasets are in accordance with Queensland University of Technology’s Low Risk Research Project Approval # 0900001080.

References

Alotaibi F and Johnson F (2020) Why we like Google Scholar: Postgraduate students’ perceptions of factors influencing their intention to use. Aslib Journal of Information Management. Epub ahead of print 28 August 2020. DOI: 10.1108/AJIM-10-2019-0304.

Association of College and Research Libraries (ACRL) (2015) Framework for Information Literacy for Higher Education. Chicago: Association of College and Research Libraries. Available at: http://wwwALA.org/acrl/standards/ilframework (accessed 19 August 2020).

Bates MJ (1979) Information search tactics. Journal of the American Society for Information Science 30(4): 205-214.

Bates MJ (1987) How to use information search tactics online. Online, 11(5): 47-54.

Bates MJ (1989) The design of browsing and berrypicking techniques for the online search interface. Online Review 13(5): 407-424.

Bates MJ (2019) How search really works. UX Radio. Available at: www.soundcloud.com/ux-radio/how-search-really-works-with-dr-marcia-bates (accessed 19 August 2020).

Berliner D (1994) Expertise: The wonder of exemplary performances. In: Mangieri JN and Block CC (eds) Creating Powerful Thinking in Teachers and Students: Diverse Perspectives. Fort Worth, TX: Harcourt Brace College, pp. 141-186.

Blachman N (2015) Google guide. Available at: www.googleguide.com/creative_commons_license.html (accessed 19 August 2020).

Bozdag E and Van den Hoven J (2015) Breaking the filter bubble: Democracy and design. Ethics and Information Technology 17: 249-265.

Brace W (2012) A Model-Based Approach to Requirements Analysis for Complex Products in Engineering Design. PhD Thesis. Aalto University, Finland.

Branch JL (2000) Investigating the information-seeking processes of adolescents: The value of using think alouds and think afters. Library and Information Science Research 22(4): 371-392.

Charmaz K (2003) Qualitative interviewing and grounded theory analysis. In: Holstein JA and Gubrium JF (eds) Inside Interviewing: Lenses, New Concerns. Thousand Oaks, CA: SAGE, pp. 311-330.

Charmaz K (2014) Constructing Grounded Theory, 2d ed. Thousand Oaks, CA: SAGE.

Chu SK and Law N (2007) Development of information search expertise: Research students' knowledge of source types. Journal of Librarianship and Information Science 39(1): 27-40.

Crotty D (2019) Access alone isn’t enough: Revisiting calls for discovery, infrastructure, technology, and training. Available from: scholarlykitchen.sspnet.org/2019/01/31/access-alone-isnt-enough-revisiting-calls-for-discovery-infrastructure-technology-and-training/ (accessed 27 August 2020).

Davis K (2014) Constructivist grounded theory. In: Bruce C, Davis K, Hughes H et al. (eds) Information Experience: Approaches to Theory & Practice. Bingley, UK: Emerald Publishing, pp. 38-39.
Dey I (1999) *Grounding Grounded Theory*. San Diego: Academic Press.

Dreyfus SE (2004) The five-stage model of adult skill acquisition. *Bulletin of Science, Technology & Society* 24(3): 177-181.

Duarte E, Oliveira Jr E, Côgo F et al. (2015) Dico: A Conceptual Model to Support the Design and Evaluation of Advanced Search Features for Exploratory Search. *Lecture Notes in Computer Science*, 9299: 87-104.

Edwards SL (2004) Web-based information searching: Understanding student experiences to enhance the development of this critical graduate attribute. In K. Appleton, C. Macpherson, & D. Orr (Eds.). *Third International Lifelong Learning Conference*, pp. 106-115.

Edwards SL (2006) *Panning for Gold: Information Literacy and the Net Lenses Model*. Adelaide, Australia: Auslib Press.

Edwards SL (2017) *Net Lenses: Understanding Student Experiences to Enhance the Learning Environment*. Available at: https://research.qut.edu.au/netlenses/ (accessed 29 August 2020).

Edwards SL and Bruce CS (2002) Needles, haystacks, filters and me: The IT confidence dilemma. *Lifelong Learning Conference*, Yeppoon, Central Queensland, Australia, 16-19 June 2002.

Edwards SL and Bruce CS (2010) The assignment that triggered change: assessment and the relational learning model for generic capabilities. *Assessment and Evaluation in Higher Education* 29(2): 141-157.

Eickhoff C, Dungs S and Tran V (2015) An eye-tracking study of query reformulation. *SIGIR’15 Proceedings*, 9-13 August 2015, Santiago, Chile, pp. 13-22.

Else.is (2018) *Sample pages*. Available at: http://else.is (accessed 19 August 2020).

Ericsson KA and Simon H (1993) *Protocol Analysis: Verbal Reports as Data*. 2nd edn. Cambridge, MA: MIT Press.

Fessl A, Wertner A and Pammer-Schindler V (2018) Digging for gold: Motivating users to explore alternative search interfaces. *Lecture Notes in Computer Science* 11082: 636-639.

Fessl A, Apaolaza A, Gledson A et al. (2019) Mirror, mirror on my search: Data-driven reflection and experimentation with search behaviour. *Lecture Notes in Computer Science* 11722: 83-97.

Fister B and MacMillan M (2020) *Information literacy in the age of algorithms*. Available at: www.projectinfolit.org/uploads/2/7/5/4/27541717/algoreport.pdf (accessed 19 August 2020).

Francis C (2004) Designers search strategies influenced by the interaction with information retrieval systems (IRS): Within the early stages of the design process. Proceedings of *APCHI*, Rotorua, New Zealand. *Lecture Notes in Computer Science* 3101: 646-650.

Ghosh S and Shah C (2017) Information seeking in learning-oriented search. *Proceedings of the Association for Information Science and Technology* 54(1): 682-684.

Ghosh S, Rath M and Shah C (Mar 1, 2018) Searching as Learning: Exploring Search Behavior and Learning Outcomes in Learning-related Tasks. *CHIIR ’18: Proceedings of the 2018 Conference on Human Information Interaction & Retrieval*, March 2018, pp. 22–31.
Google Scholar (n.d.) About. Available at: scholar.google.com/intl/en/scholar/about.html (accessed 19 August 2020).

Google Scholar (n.d.) Inclusion guidelines. Available at: scholar.google.com/intl/en/scholar/inclusion.html (accessed 19 August 2020).

Google Webmasters Support (2019). Structured data markup for datasets. Available at: support.google.com/webmasters/thread/1960710 (accessed 19 August 2020).

Halevi G, Moed H and Bar-Ilan J (2017) Suitability of Google Scholar as a source of scientific information and as a source of data for scientific evaluation: Review of the literature. Journal of Informetrics 11(3): 823-834.

Han Z, Hansen P, Xu H et al. (2020) Examining the classification and evolution of novice users’ mental models of an academic database in the search task completion process. Journal of Information Science 46(2): 205-225.

Hansen P and Rieh SY (2016) Editorial: Recent advances on searching as learning: An introduction to the special issue. Journal of Information Science 42(1): 3-6.

Hsieh-Yee I (1993) Effects of search experience and subject knowledge on the search tactics of novice and experienced searchers. Journal of the American Society for Information Science 44(3): 161-174.

Hsieh-Yee I (2001) Research on Web search behavior. Library and Information Science Research 23(2): 167-185.

Hunt T (2017) Googling Gives Us Answers—but Deprives Us of Intelligence. Available at: qz.com/961379/googles-search-algorithms-act-as-our-brains-but-what-are-they-trying-to-get-us-to-think (accessed 19 August 2020).

Hunt T (2020) What might a post-Google web search engine actually look like? Available at: medium.com/ted-hunt/what-might-a-post-google-web-search-engine-actually-look-like-bdacf539f435 (accessed 19 August 2020).

Institute of Electrical and Electronics Engineers (IEEE) (2019) Engineering and Technology Wiki: List of IEEE Milestones. Available at: ethw.org/Milestones/List_of_IEEE_Milestones (accessed 29 August 2020).

Just N and Latzer M (2017) Governance by algorithms: reality construction by algorithmic selection on the Internet. Media, Culture & Society 39(2): 238-258.

Kvale S and and Brinkmann S (2014) InterViews: Learning the Craft of Qualitative Research Interviewing. 3rd edn. London: SAGE Publications.

Latzer M, Hollnbuchner K, Just N et al. (2016) The economics of algorithmic selection on the Internet. In: Johannes M. Bauer and Michael Latzer (eds) Handbook on the Economics of the Internet. Northampton, MA: Edward Elgar Publishing, pp. 395-425.

Laubheimer P (2016) Seven ways to improve your website’s or intranet’s built-in search engine. Available at: https://www.nngroup.com/articles/internal-website-search/ (accessed 29 August 2020).
Leslie I (2014) Google makes us all dumber: The neuroscience of search engines. *Salon*, 12 October 2014. Available at: https://www.salon.com/2014/10/12/google_makes_us_all_dumber_the_neuroscience_of_search_engines/ (accessed 29 August 2020).

Lillard LL and Ha Y (2015) Bates’ Berrypicking Model (1989, 2002, 2005). In: Al-Suqri MN and Al-Aufi AS (eds) *Information Seeking Behavior and Technology Adoption: Theories and Trends*, pp. 113-123.

Luh C, Yang S and Huang TD (2016) Estimating Google’s search engine ranking function from a search engine optimization perspective. *Online Information Review* 40(2): 239-255.

Marchionini G (2019) Search, sense making and learning: closing gaps. *Information and Learning Sciences* 120(1/2): 74-86.

Marton F and Booth SA (1997) *Learning and Awareness*. New York: Psychology Press.

Marton F (1994) On the structure of awareness. In: Bowden JA and Walsh E (eds) *Phenomenographic Research: Variations in Method*, pp. 176-205.

Martzoukou K (2008) Students' attitudes towards Web search engines: Increasing appreciation of sophisticated search strategies. *Libri* 58(3): 182–201.

Meyer JHF and Land R (2003) Threshold concepts and troublesome knowledge (1): Linkages to ways of thinking and practising within the disciplines. In: Rust C (ed) *Improving Student Learning: Ten Years On*. Oxford, UK: Oxford University Press, pp. 1-16.

Moraveji N, Russell D, Bien J et al. (2011) Measuring improvement in user search performance resulting from optimal search tips. Proceedings of the 34th ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 355–363.

Ng L (2020) *Change in Google Search is killing it*. Available at: https://uxdesign.cc/change-in-google-search-is-killing-it-e52a098b271e (accessed 29 August 2020).

Nielsen J (2000) *Novice vs. expert users*. Available at: https://www.nngroup.com/articles/novice-vs-expert-users/ (accessed 29 August 2020).

Nielsen J (2004) *When search engines become answer engines*. Available at: https://www.nngroup.com/articles/search-engines-become-answer-engines/ (accessed 29 August 2020).

Nielsen J (2010) *Testing expert users*. Available at: https://www.nngroup.com/articles/testing-expert-users/ (accessed 29 August 2020).

Noble SU (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press.

Oh K and Colón-Aguirre M (2019) A comparative study of perceptions and use of Google Scholar and academic library discovery systems. *College & Research Libraries* 80(6): 876-891.

Oh KE and Tang R (2018) How do university students receive, read, find, share, and store news? A survey study on mobile news behavior. *Journal of Library & Information Studies* 16(1): 1-21.

Okhovati M, Sharifpoor E, Aazami M et al. (2017) Novice and experienced users’ search performance and satisfaction with Web of Science and Scopus. *Journal of Librarianship and Information Science* 49(4): 359-367.
Ondrusek A, Ren X and Yang C (2019) A content analysis of errors in MLIS students’ online searching assignments. *Journal of Education for Library and Information Science* 60(4): 312-328.

Pariser E (2011) *The Filter Bubble: How the New Personalized Web is Changing what we Read and how we Think*. New York: Penguin Books.

Price C (2020) *Seventeen Great Search Engines You can use Instead of Google*. *Search Engine Journal*. Available at: https://www.searchenginejournal.com/alternative-search-engines/271409 (accessed 19 August 2020).

Rieh SY, Collins-Thompson K, Hansen P 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.

Robertson SE (2016) Search then and now, with particular reference to the Web. *International Society for Knowledge Organization, Annual Tony Kent Strix Lecture*, 31 October 2016, London UK.

Rubin HJ and Rubin IS (2011) *Qualitative Interviewing: The Art of Hearing Data*. 3rd edn. Thousand Oaks, CA: SAGE Publications.

Russell DM (2019) *The Joy of Search*. Cambridge, MA, USA: MIT Press.

Salehi S, Du JT, and Ashman H (2019) Use of Web search engines and personalisation in information searching for educational purposes. *Information Research* 23(2), paper 788. Available from: http://InformationR.net/ir/23-2/paper788.html (accessed 19 August 2020).

Schön DA (1983) *The Reflective Practitioner: How Professionals Think in Practice*. New York: Basic Books.

Schön DA (1987) *Educating the Reflective Practitioner*. San Francisco: Jossey-Bass.

Schonfeld RC (2015) Meeting researchers where they start: Streamlining access to scholarly resources. *Ithaka S+R*, 26 March 2015. Available at: https://doi.org/10.18665/sr.241038 (accessed 19 August 2020).

Shieber J (2020) Google backtracks on search results design. *TechCrunch*, 24 January 2020. Available at: https://techcrunch.com/2020/01/24/google-backtracks-on-search-results-design/ (accessed 19 August 2020).

Siemens AG (2017). Semantic search tool for document tagging, indexing and search: Semantic knowledge management tool (SKMT). U.S. Patent 9684683B2, expiration 2032-03-06.

Simon HA and Chase WG (1973) Skill in chess. *American Scientist* 61: 394-403.

Srinivasan R (2018) Keynote: The people’s Internet: How to arrive at a digital world that respects human values. *Annual Conference of the Association for Information Science and Technology*, 11 November 2018, Vancouver, BC, Canada.

StatCounter (2020) *Search engine market share worldwide*. Available at: https://gs.statcounter.com/search-engine-market-share (accessed 19 August 2020).

Summit RK (2002) Reflections on the beginnings of Dialog. *Association of Independent Information Professionals (AIIP) Connections* 16(1): 11-15.

Theng Y, Lee EA, Chu SK et al. (2016) Scaffolding in information search: Effects on less experienced searchers. *Journal of Librarianship and Information Science* 48(2): 177-190.
Tucker VM (2016) Learning experiences and the liminality of expertise. In: Land R, H.F. Meyer HF and Flanagan MT (eds) *Threshold Concepts in Practice*. Rotterdam: Sense Publishers, pp. 93-106.

Tucker VM (2020) Curriculum design on the edge: A case illustration of liminal learning activities for search expertise. In: Timmermans JA and Land R (eds) *Threshold Concepts on the Edge*. Leiden, The Netherlands: Brill, pp. 239–254.

Tucker VM, Bruce C and Edwards SL (2016) Using grounded theory to discover threshold concepts in transformative learning experiences. In: Tight M and Huisman J (eds) *Theory and Method in Higher Education, Vol. 2*. Bingley, UK: Emerald Group Publishing, pp. 23-46.

Tucker VM, Weedman J, Bruce CS et al. (2014) Learning portals: Analyzing threshold concept theory for LIS education. *Journal of Education for Library and Information Science* 55(2): 150-165.

Tufekci Z (2015) Algorithmic harms beyond Facebook and Google: emergent challenges of computational agency. *Colorado Technology Law Journal* 13(2): 203-217.

Vaidhyanathan S (2012) *Googlization of Everything: And Why we should Worry*, Updated edition. Berkeley, CA: University of California Press.

Vaidhyanathan S (2018) *Antisocial media: How Facebook disconnects us and undermines democracy*. New York: Oxford University Press.

Vega LC (2017) Novices and expert users, not novices or expert users. *Society for Technical Communication Summit Proceedings*, Washington DC, 7-10 May 2017, pp. 193-197. Available at: https://www.stc.org/wp-content/uploads/2017/06/STC-Summit-Proceedings-2017.pdf (accessed 19 August 2020).

Walton G, Pickard AJ and Dodd L (2018) Information discernment, mis-information and pro-active scepticism. *Journal of Librarianship and Information Science* 50(3): 296-309.

Winograd T and Flores F (1986) *Understanding Computers and Cognition: A New Foundation for Design*. Norwood, NJ: Ablex Publishing.

Zittrain J (2019) Annals of technology: The hidden costs of automated thinking. *The New Yorker*, 23 July 2019. Available at: https://www.newyorker.com/tech/annals-of-technology/the-hidden-costs-of-automated-thinking (accessed 19 August 2020).

Author Biographies

**Virginia M. Tucker** is assistant professor at the School of Information, San José State University. Her research and teaching interests are in information retrieval system design, advanced search, and information architecture. Tucker was previously product architect and training manager at Dialog/Thomson (now ProQuest), the physics librarian at Stanford University, and a public law librarian. She has a PhD in information systems from Queensland University of Technology; MLS from University of California at Berkeley; and a BA in music from Stanford University.

**Sylvia L. Edwards**, Dip.Lib.(RMIT), GCEd(HE), MIT(Res.), PhD(QUT), AALIA, MACS(senior), is a Professor in Information Systems at Queensland University of Technology (QUT), Faculty of Science. She researches higher education leadership, information searching behaviour and information literacy, specialising in applying her research findings in practice. Her publications are available at http://eprints.qut.edu.au/view/person/Edwards_Sylvia.html.
Appendix A. Net Lenses Original Model: Four categories of web search experiences

The four categories of the original Net Lenses model (Edwards, 2006: 2017) are summarised below, according to the four parameters in the model’s characterisation of the search experience: the searcher’s awareness structures, approach to learning, response to obstacles, and search outcomes.

| Net Lenses Model: Searcher Experience | Category 1 | Category 2 | Category 3 | Category 4 |
|--------------------------------------|------------|------------|------------|------------|
| General characterisation:            | “looking for needle in haystack” | “finding way through a maze” | “using the tools as a filter” | “panning for gold” |
| Parameter 1: Awareness structures (searcher’s focus) | Searcher’s worldview | The topic | The topic and tools | Focus is the tool’s structure and the topic is secondary | The topic, the structure of the tool, and quality of the information resource |
| Information environment | aware of information environment but not the structure or its importance | interest in and awareness of structure of information environment is beginning | strong awareness | strong awareness |
| Tool structures | uses tools but not aware of details of tools or their structures | willing to try advanced search tools | strong awareness | strong awareness and use of structure of tool |
| Information quality | not of interest and not understood | aware of information quality but it is not a major focus | aware that information quality is important but it’s not a real factor in searching | strong awareness; interested in primary, rather than secondary sources |
| Parameter 2: Approach to learning | Confidence with IT generally | moderate confidence with IT; confidence may interfere with appreciation of structure of environment | good confidence with IT but still doesn’t notice own mistakes | high confidence with IT; more aware of mistakes and likely to self-correct | high; aware of possible mistakes; more likely to self-correct, ask for help |
| Use of search planning | poor or non-existent | has commenced planning process | planning is evident; may even write down terms to use and consider synonyms | planning is evident and may be written down before searching; often analyses term and synonyms before searching |
| Parameter 3: Response to obstacles | Persistence level | low; likely to abandon search | moderate; likely to consider alternatives and persevere | high | very high |
| Assumptions made | views problems as due to weakness in the tool or content of database(s) | still blames tool rather than question own abilities | looks to self for cause of mistakes and corrects them | looks to self for cause of failed search, analyses, and re-attempts |
|------------------|------------------------------------------------------------------------|-----------------------------------------------------|---------------------------------------------------|---------------------------------------------------------------|
| Parameter 4: Search outcome | search is likely to be abandoned; outcome has been affected by poor planning and lack of reflection | search may be successful | search is usually successful; may miss things as reflection is not planned | search is successful |
Appendix B. Study Participants, Search Tasks, and Interview Questions

Study Participants. The 20 participants included:

- 9 students in an MLIS degree programme who had been pre-selected by one of two instructors of the advanced online searching course for having demonstrated expertlike abilities.
- 11 information professionals for whom advanced search methods was a core part of their position. They included LIS faculty who teach advanced search, information brokers, and search engine developers. They averaged 32.7 years of relevant professional experience.

Search Tasks. Each participant conducted a known item search (where the objective is to locate a specific answer or document), then a subject search (where exploratory search tactics were needed, but no domain-specific knowledge was required, e.g., no medical or legal terminology). Think aloud protocols (Ericsson and Simon, 1993) were used, that is, the participant was prompted to narrate their thoughts and reasons for decisions made while conducting the searches.

Interview Questions. Questions were posed both before and after the search tasks.

Before the search tasks were, the following questions about search experiences were asked:

- What experiences helped you learn to search? What learning experiences do you think helped to make you the searcher you are today?
- Please describe one search experience in the past and what you learned from it.
- In your opinion, what are the key concepts you need to learn to be an expert searcher?
- Can you remember learning any specific concepts that allowed you to search in a more expert way?

After the search tasks, the interview questions (talk-after protocol) included:

- In your opinion, what are the key concepts you need to learn to be an expert searcher? Can you remember learning any concepts that allowed you to search in a more expert way? (These questions were revisited because the activity of searching can trigger further recollections.)
- Do you have any writings that document your learning experiences as a searcher? Would you be willing to share these? (If the participant had writings to share, this was handled later.)

Research protocols for conducting semi-structured interviews were used, such as probing, indirect, and follow-up questions, as well as well-timed pauses for the participant to reflect and associate their thoughts (Kvale and Brinkman, 2014; Rubin and Rubin, 2011).