

Making the Most of Online Searches

April 01, 2011



APS Members are editing and updating entries on psychological concepts, methods, and people on Wikipedia as part of the APS Wikipedia Initiative (APSWI). I hope you will join those efforts to ensure that such entries are clear, accurate, and accessible. To help make the most of online searches, the search engine Google has a special team dedicated to Search Quality and User Happiness. In this month's column, a member of that team offers insights into making the most out of Internet search results.

-Mahzarin R. Banaji

APS President

On the face of it, you'd think that searching on a modern search engine such as Google is a pretty simple and straightforward skill. And mostly, you'd be right. It's the exceptions that are interesting.

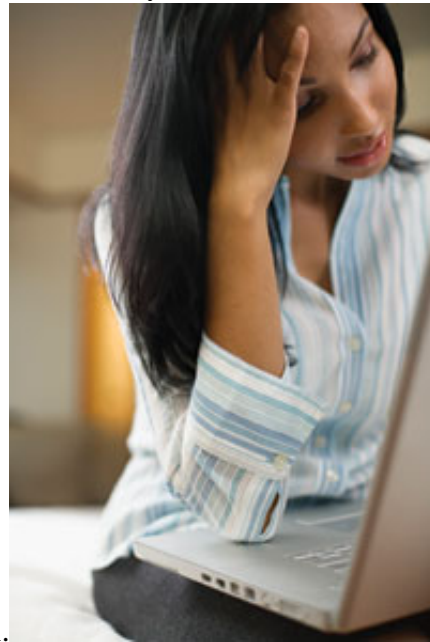
We've all had the experience of trying to find something that we just can't quite seem to nail down. You expect there's a Web page out there somewhere in the billions of possible pages, but you just can't figure out what to do to make it pop to the top of the search results. Even more strangely, you've probably also been frustrated trying to find something, then told a colleague about it, only to have the perfect result show up when *they* did what turned out to be exactly the right query.

That's where the skill of search comes into play. Sometimes you (or your colleague) do just get lucky and manage to phrase the query just so. The good news is that search engines do a remarkably good job on the vast majority of searches. More than almost any other technology, search engines have transformed the way we do research: Papers and results that were previously undiscoverable (or only painfully and laboriously discoverable) now have become simple and quick to locate. Where graduate students once slaved endlessly over a hot photocopier in library stacks, they can now run a search and locate the relevant papers, as well as all the papers those papers reference, and so on. The scholarship of research really is different these days. It's not just simpler, but also broader and deeper. Given the same

number of research hours, one can potentially reach much, much more.

Back to that sticky search problem: What causes difficult search tasks, and what can we do to work around the less obvious search problems?

These problems are often called “long tail” problems because they’re tasks that take more than the average number of searches to accomplish. As with most internet-related human behaviors, search tasks follow a power distribution curve in the number of searches needed to satisfy a search goal. If you’re just looking for the main Web site of a university department, that’s pretty simple to do and takes only a search or two. But if you’re trying to understand the latest research findings on the early detection and treatment recommendations of autism, that’s a task that will take many searches over an extended period



of time. Long tasks are well out on the long part of the tail.

Difficult search tasks are difficult for a number of reasons. First and foremost is an effect well known in cognitive psychology: the framing effect. When a searcher initially conceives of a search task, the problem is sometimes framed in terms that are relevant to the searcher, but not necessarily in the language of the literature. This most commonly happens when the searcher seeks out information in a domain in which they are not expert. A friend recently spent a great deal of time searching for the data set on which gestational diabetes blood sugar levels were established, but she had no success until she had determined that the appropriate word to use in searching was “pre-prandial” (as in “pre-prandial blood glucose levels”). Once she discovered that word, the world of scholarly literature about pre-prandial glucose testing was easily found; without the key term, it was merely a long slog through general results. In this case, the language she used to frame the question predisposed the search engine to the non-technical literature. Once a key term for a search task is found, an entirely new universe of results suddenly becomes open for inspection.

Luckily, this “key term effect” happens primarily in technical domains where the long-tail effect dominates on relatively rare searches. In more common search tasks, many other people have followed common search paths that led to a successful outcome. In these tasks, search engine automatic synonymization works very effectively to get the searcher to their results rapidly. Although Google will

synonymize “blood sugar” with “glucose,” too few people search for pre-prandial tests norms for the term “pre-prandial glucose” to be suggested as a synonym for “blood sugar test.”

Another method for getting out of a framing mindset is to check out the related searches that other searchers have used. (Related searches are shown in the left hand navigation panel or at the top of the organic results.) These searches made by others working in the same domain can often lead to useful re-framings of the search task. The query [power law] might have other related searches such as [pareto], [power distribution], [Zipf distribution], or [80 20 rule]—all useful suggestions that might easily unstuck a conceptual fixedness.

Like many expert behaviors, search is one that rewards skill development, practice, and attention. In our studies, we have shown that spending a modest amount of time learning the attendant skills of search pays off in much reduced search times and search accuracy. For example, a crucial skill in reading a search-results page (indeed, any online document) is knowing how to *Find* a word on the page. In all browsers, the Control-F/Command-F/ Edit>Find function is used to locate a given word in the document. Surprisingly, our surveys show that roughly 90 percent of the U.S. English-speaking population does NOT know this key skill. Once told about this, many search tasks are significantly simplified, and on long documents (e.g., that 150-page PDF highly technical monograph you’re reading) the task of locating relevant information goes from extremely difficult to trivial.



Similarly, the skill of *scoping* a search by limiting searches to a particular resource can be very useful. Some difficult searches suffer from having search terms that are too common, making it difficult for the searcher to separate the wheat from the chaff. For instance, suppose you have a vague recollection that there was an interesting article about crossword puzzles in some issue of the *APS Observer*. Doing a general Web search for [crossword] is unlikely to bring any issue of the *Observer* into the top 10 results. Knowing how to limit your search to just the *Observer* website will bring all of the articles about crossword puzzles into immediate focus. The way to do this *scoped search* is by using the “site” operator — for example: [site:psychologicalscience.org crossword]. This is a handy skill to have when you want to search different repositories. (And of course, if your search comes up empty, be sure to try your search without the “site” operator.)

But the surgical scalpel of key word choice cuts both ways, sometimes pushing your search investigation into a particular context that you might not have considered. For example, when searching for a door lever to be put onto a child’s door handle to simplify door opening, including the search term “child” automatically puts your searches into the realm of children. If you’re searching for a typically childhood disease or disorder, that’s usually good, but in the context of doorknobs and handle extensions, queries like [child door knob] are dominated by the much more typical problem of *preventing* children from opening particular doors. For this particular search, rethinking the problem in terms of other uses leads to a better search experience. For instance, older people also have problems opening doors because of

reduced grip strength and imprecise motor movements. The search [elder door knob] gives much better results for this particular search task of looking for aids for door opening. The important thing to remember is that Web search operates over the entire Web, which might well include topics and areas you might not be considering when you form an initial concept for the search.

Of course, a key thing to remember about Web search is that both the contents of the Web and what search engines can do to process that content are constantly under revision. What this means for you is that search is a skill like any other. It is useful for professionals to pay attention to new content resources as they come online (that is, as they become accessible through search engines), watching for new search capabilities (such as the ability to search realtime streams for breaking news on current events), and new ways of viewing the results of searches (timeline views of search results). There are classes, information streams, and resources available for staying on top of what's going on. To be the best possible searcher, you need to make time to track these new capabilities as well as understand the entailments of what's possible.

With the inexorable and rapid transformation of the World Wide Web into a resource of incredible depth and breadth, you owe it to yourself to stay in touch as new materials and new tools transform research problems from very difficult or impossible to quick and simple tasks. The Web changes, and so do the tools; keep learning about what's possible. æ