Librarians' Articles   Librarians' Publications

2011

Relevance, Choices, and the Goldilocks Problem

Mary Whisner
University of Washington School of Law

Practicing Reference . . .

Relevance, Choices, and the Goldilocks Problem*

Mary Whisner**

Ms. Whisner ponders a core question in answering reference queries: How can we know whether what we find is relevant to what the questioner wants? Her article provides criteria to consider and some guidelines for choosing sources in response to a query.

¶1 Early in the school year, one of our reference interns asked how she should decide whether an article was relevant to a professor's question. I joked that if she could figure that out, she could skip past the master's program and go right for a Ph.D. in information science, writing a dissertation on "aboutness." More seriously, we talked about specifics of the question. The professor had asked for articles discussing a particular aspect of humanitarian assistance. The intern hadn't found entire articles on the topic, but had found some articles that discussed it for a couple of paragraphs or a section. I suggested that she send those articles to the professor with an explanation in her cover memo about why she was sending them.

¶2 This problem of selecting sources comes up whether researchers are doing work for themselves or (like the intern) for others. It is challenging because it involves concurrent judgments in different dimensions. There's the intern's question: Is this article relevant? But there are also questions of quantity (do we have enough? too much?), quality (is it reliable? authoritative? current?), and personal preferences of the requester.

What's Relevant?

¶3 Not long ago, an LL.M. student came to the reference office and asked for help using Westlaw. She was doing an assignment that included a long hypothetical about financial transactions among various parties, followed by a question that asked for two relevant cases discussing whether a bankruptcy creditor could pursue certain assets. The textbook hinted that it would be easy to find the cases by first finding a relevant law review article. The student had typed some words into a natural language search, which of course returned one hundred articles, but she couldn't see how to get from there to the answer.

¶4 After a long, confused conversation, I came to see the underlying problem: she had not studied the hypothetical closely enough to put the right words into her search or to recognize a relevant article when it showed up in the search results. At base, she did not have a good sense of the question to ask. I did not think it would help her education for me to state the question or construct a search or say, "Here, this article is relevant; look at the cases it discusses." A big part of her learning experience should be wrestling with the hypothetical. So I encouraged her to reread it and to talk it over with her classmates so that she could try to state a better question and come up with search terms.

¶5 In the courtroom, "'[r]elevant evidence' means evidence having any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence."[1] A piece of evidence does not have a permanent label saying it's relevant or it's not. Relevancy "exists only as a relation between an item of evidence and a matter properly provable in the case. Does the item of evidence tend to prove the matter sought to be proved?"[2] In research, relevant sources are those that tend to answer the question you're working on. The LL.M. student's natural language search retrieved articles that matched the words she'd typed in (according to the proprietary algorithm that approximates relevance), but because she did not have a clear question, she couldn't spot which of the articles could help her figure out how to advise the hypothetical client. No question, no relevant sources.

¶6 Sometimes we need to work with patrons to help them form questions from the whirl of facts and problems they face. For instance, this unhappy statement: "My ex-husband moved out of state and now expects me to drive the kids to see him like I used to do when we lived in the same town. I just don't have the time to be driving seven hours every Saturday. He can't really expect me to do this, can he?" might morph into: "How can I get our parenting plan modified because of changed circumstances?" Reformulating the question makes it possible to find relevant materials to address it.

¶7 Reference librarians often are presented with well-formed questions: Are there any Washington cases discussing intentional infliction of emotional distress when the plaintiff was a bystander to the accident? Could you help me find recent law review articles on humanitarian aid after natural disasters? How many dialysis centers are there in the United States? How many kidney transplants are performed annually? What studies have looked at racial disparities in Washington's criminal justice system? We might see ambiguities in some questions, but at least we have something fairly specific in front of us. We can ask for clarification if we need it or make our best guess and tell the requester our assumptions.

¶8 But having the question handed to us only saves us the first of several steps of analysis: (1) figuring out the question; (2) picking a database (or print source) that is likely to have relevant documents; (3) constructing a search (or choosing an index term) that's likely to retrieve relevant documents; (4) skimming the docu-

---

1. FED. R. EVID. 401.
2. *Id.* comment.

ments to judge which are relevant; and (5) deciding which documents to use. I have seen people struggle—and I've struggled myself—at each stage.

¶9 Picking a database[3] and constructing a search involve informed guesswork. What would a relevant document be? What would it look like? What words would it have in it, and in what relation? The next step after running a search—the process of determining which documents are relevant—requires fresh judgment. If I feel I was awfully clever constructing the search, then it's tempting to think that almost everything the search retrieved is relevant—and that no documents were missed by the search. But there are often surprises. What seemed like a terrific search—great terms, appropriate synonyms, perfect subject headings or key numbers—can be a bust, retrieving nothing or a lot of irrelevant documents. It's important to look at the documents and assess not whether they satisfy the search (of course they will) but whether they actually answer the original question. When I go fishing in article databases, I often find myself skimming dozens of articles and abstracts for every few that I select. And it's common to have to go back a step or two, trying a different database or search. My first effort might not have been so awfully clever after all. Even if it was clever, it didn't do the job.

Which Source to Use

¶10 Sometimes it's clear when a source answers a question. For instance, if I want to know the date that the jury returned its verdict in the punitive damages phase of the *Exxon Valdez* trial, here is a good source:

> Source: Joint Opening Brief of Appellants Exxon Corporation and Exxon Shipping Company at 7–8, In re: Exxon Valdez, 239 F.3d 985 (9th Cir. 2001) (No. 99-35898), 1999 WL 33622384.

> Answer: "The class punitive claims were then tried in district court in a three-phased trial that commenced on May 2, 1994 and continued until September 16, 1994, when the jury returned a verdict awarding the class the enormous sum of $5 billion in punitive damages."

This brief seems like a credible source: the legal team working on this multi-billion-dollar case must have had access to the trial record and also had a strong incentive to proofread carefully.

¶11 But a relevant, reliable source is not all a researcher wants. For some reason or other, I might prefer not to cite Exxon's brief. Maybe I don't want to appear partisan by using one side's brief (or maybe I am partisan and don't want to use Exxon's brief because of my bias).

¶12 Perhaps it would look more scholarly to cite a law review article:

> Source: William H. Rodgers, Jr., *Growth and Form: Indian Tribes, Terrorism, and the Durability of Environmental Law*, 26 VT. L. REV. 865, 869 n.21 (2002).

> Answer: "*See* In re Exxon Valdez, 229 F.3d 790, 794 (9th Cir. 2000) ("On September 16, 1994, the jury awarded punitive damages in the sum of $5 billion, at that time the largest award of its kind in history.")."

---

3. WestlawNext reduces the importance of this step—one can search across many databases—but one would still need to choose WestlawNext over, say, Google or a print source. And it is almost always more efficient to exercise some control over the databases searched.

However, instead of citing this law review article, I would probably prefer to go back a step and use the case that it cites.

¶13 But if I'm writing for readers who are not used to finding cases, I might want to cite something that's easier to locate. And if I'm writing for the web, I might want something that I can link to that doesn't require a password. In Google Books I find:

> Source: Mark Hansen, *Record Punitives: $5 Billion Award Unlikely to Stand*, A.B.A. J., Dec. 1994, at 41, *available at* http://books.google.com/books?id=MSo7JKWXiFkC&lpg=PP1&lr &rview=1&pg=PA41#v=onepage&q&f=false.
>
> Answer: "The Sept. 16 verdict, which capped a 20-week trial over the nation's worst oil spill, was the largest punitive damage award ever."

¶14 The *ABA Journal* is a good, mainstream source. But I notice that the article about the September 16 verdict didn't appear until the December issue. So I decide to look for contemporaneous coverage in the *New York Times*.

> Source: Keith Schneider, *Exxon Is Ordered to Pay $5 Billion for Alaska Spill*, N.Y. TIMES, Sept. 17, 1994, at 11, *available at* http://www.nytimes.com/1994/06/14/us/jury-finds-exxon -acted-recklessly-in-valdez-oil-spill.html.
>
> Answer: "A Federal jury in Anchorage yesterday ordered the Exxon Corporation to pay $5 billion in punitive damages to about 34,000 fishermen and other Alaskans who said they were harmed by the Exxon Valdez oil spill more than five years ago."

This source might satisfy all of my wishes—reliable, respected, and easy for my readers to find.[4]

¶15 If the only goal is to learn the date, you could rely on any of these sources. It was only subjective preferences about what to cite that made me bother to look at the others. In day-to-day life, these preferences are generally unspoken—we search the *New York Times* rather than the *Washington Post* or *Moore's Federal Practice* rather than *Federal Practice and Procedure*, just because we do—but preferences are there nonetheless. When we research for other people, we make judgments—explicitly or implicitly—about their preferences. If a professor asks me the date of the *Exxon Valdez* verdict, I won't give him all the newspaper and magazine articles, briefs, cases, and books that have that piece of information. I will choose one or two, perhaps using the same considerations I discussed above.

¶16 Relevance isn't everything. In a trial, the judge may exclude relevant evidence "if its probative value is substantially outweighed by . . . considerations of undue delay, waste of time, or needless presentation of cumulative evidence."[5] And when we are conducting research, whether for ourselves or others, we likewise may stop because of these considerations. Alas, there is not a handy meter that lights up when our piling up of sources becomes a waste of time or amounts to needless presentation of cumulative evidence. Again, we need to develop and use our judgment.

---

4. A happy bonus is that the *New York Times* article is just one in a collection of articles about the Exxon Valdez oil spill. *Article Collections: Valdez Oil Spill*, N.Y. TIMES, http://www.nytimes.com /keyword/valdez-oil-spill (last visited Feb. 15, 2011).

5. FED. R. EVID. 403.

The Goldilocks Issue

¶17 We all have a little Goldilocks in us. A law student working on a law review comment wants there to be some articles on the issue (so she'll have something to work with) but not too many (so her piece would be preempted). A lawyer trying to figure out a legal problem wants to find some good explanation (say, a chapter in a deskbook or a section in a treatise) but not too much (an entire treatise) or too little (a one-page bar journal piece). An undergraduate writing a term paper wants sources that are scholarly enough to help with his analysis (and satisfy the professor) but not so long or dense that they're unreadable. We all want results that are just right.

¶18 But the hoped for sources are not always there. The able researcher needs to think about how to reshape the question to get something to work with. If your wish for a Washington case that concerns a very specific fact pattern doesn't come true, then what? Look for Washington cases whose fact patterns might be analogized to yours, see if there are cases from other jurisdictions with similar facts, or look for secondary sources discussing the situation. If you really wanted a two-page overview article summarizing a complex area of law but can't find it, you might need to resign yourself to skimming longer works. Or, as in the example I opened with, if the professor asks for several law review articles on a narrow topic and all that can be found are sections of articles, then sections will have to do. If she hadn't found the porridge that was just right, an enterprising girl like Goldilocks could have made do, heating up the porridge that was too cold or cooling the porridge that was too hot.