The Relevance of Results Generated by Human Indexing and Computer Algorithms: A Study of West's Headnotes and Key Numbers and LexisNexis's Headnotes and Topics[*]

Susan Nevelow Mart[**]

This article begins the investigation into the different ways results are generated in West's "Custom Digest" and in LexisNexis's "Search by Topic or Headnote" and by KeyCite and Shepard's. The author took ten pairs of matching headnotes from important federal and California cases and reviewed the results sets generated by each classification and citator system for relevance. The differences in the results sets for classification systems and for citator systems raise interesting issues about the efficiency and comprehensiveness of any one system, and the need to adjust research strategies accordingly.

Search Modes and Search Strategies

¶1 There are two basic search modes available in today's legal research environment: those that utilize human-generated finding aids such as indexes, tables of contents, and subject headings; and those that utilize computer-generated search results, such as full-text searching using Boolean or natural language search strategies. These modes of searching do not necessarily correlate to print and online environments, because online environments proliferate with human-generated offerings: catalogs, indexes, West key numbers, and tables of contents. Each search mode offers very different benefits and drawbacks, and West key numbers and headnotes and LexisNexis's headnotes provide useful platforms for comparing the results of human-generated and computer-generated searches.[1]

¶2 Research projects generally have two main parts. The first involves understanding the general contours of a specific area of law—what Stuart Sutton calls "the development of an appropriate base-level mental model of the law under consideration . . . ."[2] Recourse to secondary sources is the general prescription for

---

    1. See *infra* ¶¶ 5–6 for a discussion of how the headnotes on the two systems are generated.

    2. Stuart Allen Sutton, Managing Legal Information: A Model of Institutional Memory Based on User Cognitive Maps 10 (Mar. 11, 1991) (unpublished Ph.D. dissertation, University of California

filling this information need. Law firm librarians rank the use of treatises, practice guides, and encyclopedias very high as a necessary research skill for new associates.[3] Young attorneys also self-report using treatises and practice guides for researching unfamiliar areas of the law.[4] And as Bob Berring has noted, what you need is "context, context, context."[5] Using a secondary source, whether it is a treatise, a practice guide, an encyclopedia, or an A.L.R. annotation, will give a researcher a framework for the legal problem and citations to the major cases (and, where applicable, the relevant statutes and regulations).

¶3 Once you have found the major cases, though, it is rare that the facts of those cases are so close to the facts of the client's case that your research is complete. The second part of the research project then begins—the search for case-specific relevant authority. The researcher needs to find other cases, similar in legal conclusions and more similar factually to the client's case. This search for more specifically relevant primary law can be called "level two research."[6] The researcher uses the major and controlling cases in the relevant area of the law (however located) as seed documents to link forward through headnotes, key numbers, KeyCite, and Shepard's or backward through headnotes, key numbers, and the cases cited in the seed cases.[7] This type of forward and backward searching from seed documents is instrumental for finding "application cases"—"cases that have only marginal value as support for an abstract proposition of law, [but] have great value in their application of the proposition to facts similar or analogous to the facts of your own case."[8]

¶4 Finding relevant cases can be a daunting task, even with good seed cases. There are so many cases. Westlaw currently has approximately nine million cases[9]

---

at Berkeley) (on file with Hastings College of the Law Library). Sutton also calls this "level-one research." *Id.*

3. Patrick Meyer, *Law Firm Legal Research Requirements for New Attorneys*, 101 LAW LIBR. J. 297, 307, 314, 2009 LAW LIBR. J. 17 ¶¶ 29, 51–52. Doing "cost-effective research" is frequently ranked number one, but in working with a roundtable of firm and academic law librarians in San Francisco to prepare a program on effective short training sessions for new summer associates, I learned that cost-effective usually means first knowing that you should begin your research in secondary sources. Cost-effective search techniques on LexisNexis and Westlaw are a separate category of cost-effectiveness.

4. Judith Lihosit, *Research in the Wild: CALR and the Role of Informal Apprenticeship in Attorney Training.* 101 LAW LIBR. J. 157, 170, 2009 LAW LIBR. J. 10 ¶ 35. ("All the attorneys I interviewed stated that if they were researching an unfamiliar area of law, they would start . . . with an appropriate secondary source, such as a practice guide, a legal treatise, or an encyclopedia, or a document repository, in order to become familiar with what one termed the 'legal landscape' . . . ." (footnote omitted)).

5. ROBERT C. BERRING, LEGAL RESEARCH FOR THE 21ST CENTURY (2000) (set of 5 videotapes) (Tape 5: *Legal Research on the Internet and Research Strategies*).

6. *See* Sutton, *supra* note 2, at 11–12. This is the part of the research process where the researcher hopes to find cases whose facts are "sufficiently similar to those of the client that a reasonable argument can be made that the court *should* similarly rule." *Id.* at 12.

7. This is often called "citation pearl growing" in the information science literature. Donald T. Hawkins & Robert Wagers, *Online Bibliographic Search Strategy Development,* ONLINE, May, 1982, at 12, 13. Because of the early development of citation systems in the retrieval of case law, this method is very highly developed for legal information. *See* Eugene Garfield, *Citation Indexes for Science: A New Dimension in Documentation Through Association of Ideas,* 122 SCIENCE 108, 108 (1955).

8. Douglas K. Norman, *The Art of Selecting Cases to Cite,* 63 TEX. B.J. 340, 340 (2000).

9. E-mail from Mark Cygnet, Thomson/West, to author (Mar. 11, 2009) (on file with author).

and LexisNexis reports a similar number.[10] Lawyers have been complaining about the ever-increasing number of cases for more than a century, and the problem has fueled any number of attempted solutions.[11] Currently, both Westlaw and LexisNexis offer digesting and citation-checking systems as methods of retrieving targeted application cases, although the systems were originally created in very different ways.

Headnote and Topic Creation[12]

¶5 LexisNexis and West Group editors create headnotes and link them to topics differently. In the West system, editors take the legal concepts from a case, summarize the concept in the editor's own language,[13] and link the resulting headnote with the appropriate key number in the West Digest System.[14] West's digest's are "basically compiled subject arrangements" of the West headnotes.[15] There is a direct correlation between the headnote and the related key number. The subject-based hierarchies of the West Digest System have been evolving since the late-nineteenth century;[16] a small portion of the outline of the over 400 topics online can be seen in figure 1.

¶6 In the LexisNexis system, the fundamental legal points of a case are "drawn directly from the court's language."[17] Then the LexisNexis headnotes have to be linked to Lexis Topics to create a linked classification system. The topics that are used to classify LexisNexis headnotes are the same topics that appear in "Search by Topic or Headnote."[18] I will refer to "Search by Topic or Headnote" as "Lexis Topics"

---

10. E-mail from LexisNexis U.S. Legal Markets Corporate Communication to author (Mar. 3, 2010) (on file with author) (estimating that LexisNexis contains 9.7 million cases).

11. In the mid-nineteenth century, the concept of citators was a response to the growing body of case law. *See* Pattie Ogden, *Mastering The Lawless Science of Our Law: A Story of Legal Citation Indexes*, 85 Law Libr. J. 1, 12 (1993). The earliest online legal databases, dating from the 1950s and 1960s, were similarly attempts to address the problems caused by the growing number of cases. *See* Jon Bing & Trygve Harvold, Legal Decisions and Information Systems 60 (1977).

12. This section is based in part on Susan Nevelow Mart, *Cite Checking: A Brave New World*, Legal Info. Alert, Apr. 2006, at 1.

13. The "editor's own language" does sometimes parallel the exact language of the court (*see, e.g.*, Tarasoff v. Regents of the Univ. of Calif., 551 P.2d 334, 335, 342 (Cal. 1976) (headnote 4)), but West editors are free to, and do, summarize legal concepts in their own words.

14. Morris L. Cohen, Robert C. Berring, & Kent C. Olson, How to Find the Law 84 (9th ed. 1989). Although the creation of headnotes is human-generated at this time, the process of matching the headnote of a target case with the language of citing cases is performed by computer algorithms. Elizabeth M. McKenzie, *Comparing KeyCite with Shepard's Online*, Legal Reference Services Q., 1999, no. 3, at 85, 90, 97 (stating that KeyCite uses automation or "computer programs" for headnote assignment).

15. Cohen, Berring & Olson, *supra* note 14, at 84.

16. *Id.*

17. Deborah Hackerson, *Legal Research: A Guide to Online Tutorials for First Year Law Students*, Legal Reference Services Q., 2006, no. 2, at 153, 159.

18. E-mail from Debra Myers, Account Executive, LexisNexis to author (Apr. 3, 2006) (on file with author). When this e-mail was sent, "Search by Topic or Headnote" was called Search Advisor. *See Lexis Search Advisor Evolves; Search Across Multiple Content Types Simulta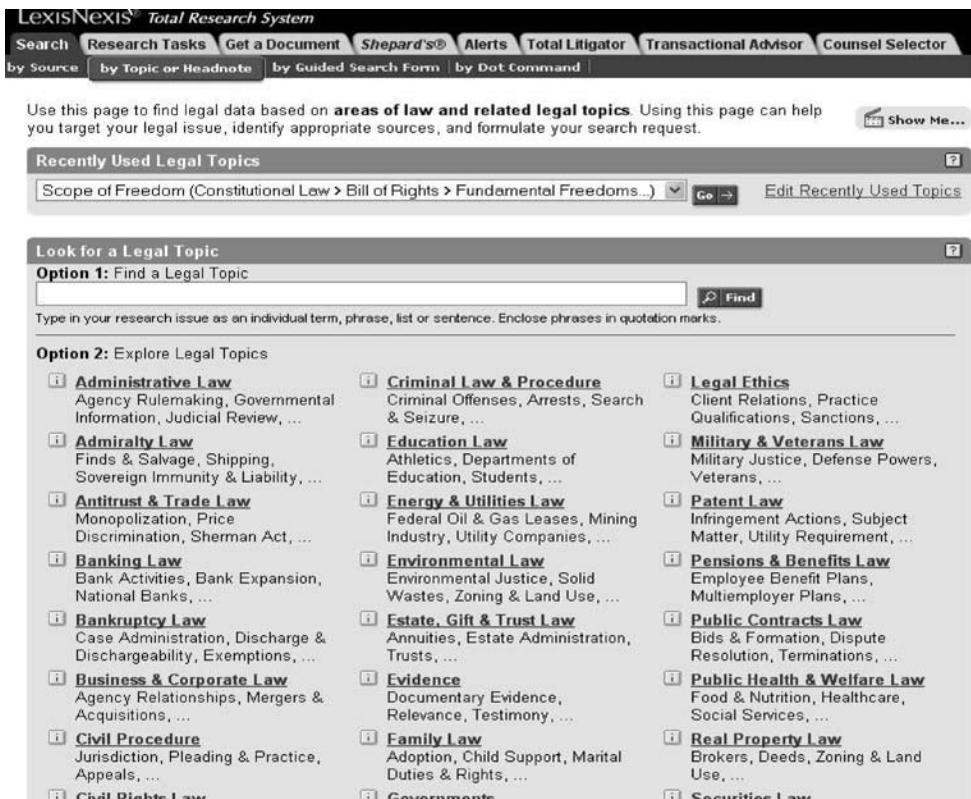neously*, LexisNexis InfoPro, Oct. 2007, *available at* http://law.lexisnexis.com/infopro/Keeping-Current/LexisNexis-Information

Figure 1. Portion of West Key Number Digest on Westlaw

Reprinted from Westlaw with permission. © 2010.

for ease of reference. Lexis Topics are the correlate for the West Key Number System, and the approximately 16,000 topics[19] represent the outline of the law as LexisNexis sees it, a portion of which is shown in figure 2.

¶7 Lexis Topics seem to be based on a patented system "for classifying concepts (such as legal concepts, including points of law from court opinions) according to a topic scheme (such as a hierarchical legal topic classification scheme.)"[20] Human

---

-Professional-Update-Newslet///Whats-New/Lexis-Search-Advisor-Evolves-Search-Across-Multiple
-Content-Types-Simultaneously/archive10-2007.

    19. E-mail from Debra Myers, *supra* note 18.

    20. Michael Ginsborg, Does Search Advisor Depend Too Little on Classifiers, and Too Much
on Algorithms, for Headnote Classification? Evidence on the Perils of Search Advisor's Automation
and the Virtue of Thomson-West Classifiers 12–13 (Aug. 2007), http://ssrn.com/abstract=1345465
(quoting System and Method for Classifying Legal Concepts Using Legal Topic Scheme, U.S.
Patent 6,502,081 (issued December 31, 2002)). Ginsborg identifies the other LexisNexis patents as:
"Computer-based System for Classifying Documents Into a Hierarchy and Linking the Classifications
to the Hierarchy," U.S. Patent 5,794,236 (issued August 11, 1998) . . . (describing algorithmic classifi-
cation of cases by means of matching case citations with topically-linked citations); and "Landmark
Case Identification System and Method," U.S. Patent Application 20060041608 (February 23, 2006)
(Search Advisor samples used to illustrate a system of cases arranged by topic in digest format, with
a user option to rank cases by frequency of citation). *Id.* at 13 n.14. LexisNexis did not acknowledge
or deny that it uses these patents to create Lexis Topics. *Id.* at 13.

Figure 2. Portion of Topic Outline on LexisNexis

editors appear to start the process by classifying sample headnotes that are part of a designated knowledge base; algorithms then take over. Algorithms extract "features"—nonstop words, phrases, and citations—from each sample headnote, ranking each feature for relevance to the assigned topic(s).[21] Classifying headnotes thus involves algorithmic assignment of topics to initially unclassified, or "candidate," headnotes, based on similarity between the candidate features and topically-ranked features of the sample headnotes. Then the features of newly classified headnotes receive topical-relevance scores, and these features are added to the knowledge base, providing further means, or feedback, for comparison between classified headnotes and candidate headnotes.[22] Thus, the role of human editors in classifying individual headnotes for each new case seems to be limited in LexisNexis.[23]

---

21. *Id.* at 16–19.

22. *Id.* at 19. This is an extreme simplification of a process described in great detail by Ginsborg.

23. *Id. See id.* at 51–53.

¶8 There are therefore two very different systems for classification. West creates a direct correlation between a headnote (drafted by a human editor) and the related key number topic, relying primarily, but not exclusively, on human editing to assign headnotes to a point in a classification system. LexisNexis relies primarily, although not exclusively, on algorithms to assign a headnote (taken from the court's language) to a topic in the classification scheme, so there is no direct correlation between the headnotes and the Lexis Topics.

¶9 If we look at the relevance of the cases that are found when using the West classification system (key numbers) and at the relevance of cases that are found using the LexisNexis classification system (Lexis Topics or "More Like This Headnote"), we should be able to see whether the difference in the method used to assign headnotes to a classification system results in a difference in the relevance and completeness of results.

The Concept of Relevance in Testing Results

¶10 Most testing of relevance in legal databases has been based on objective standards of relevance—whether the cases found do or do not fit within the particular legal concept—but in fact, each user's search needs require a unique and even shifting definition of relevance.[24] Because of the expanding universe of case law, West key numbers or Lexis Topics alone may be too broad, or may return too many results, regardless of whether the standard of relevance is objective or subjective. To be both relevant to a specific search and manageable, the researcher may need to limit the results of a broad search. And whether the digest system or citator is returning results based on human or computer classification, there are advantages and disadvantages to each system.

¶11 The difficulties inherent in relying on complex human classification systems for retrieving results are well known. In 1970, Stephen M. Marx wrote:

> Compilations of legal cases according to code numbers have been available since at least 1888. The most popular of these systems in use today is the West Key Number system. Any legal researcher will attest to the difficulty of using the *West General Digest.* The system is incredibly complex. There are, for example, at least four hundred (400) major classification headings, each with from sixteen (16) to two thousand (2000) subheadings. . . .
>
> . . . .
>
> There are four important drawbacks to systems based mainly on the use of key words and phrases: (1) these systems are static in their terminology and not adaptable to vocabulary changes; (2) these systems require that the user's thinking conform to the classifications formulated by the system designers; (3) these systems classify the law according to a rigid key word terminology without indicating the context in which the words appear;

---

24. Stuart A. Sutton, *The Role of Attorney Mental Models of Law in Case Relevance Determinations: An Exploratory Analysis,* 45 J. Am. Soc'y Info. Sci.186, 196–97 (1994). Looking at the results of a broad search, such as all cases found in the chosen *American Law Reports* annotation or in the chosen key number, may return objectively relevant results, but return as well many cases that, although perhaps relevant "topically," will not help an attorney fill in a mental map of an area of law necessitated by a particular client's situation. *Id.* "[R]elevance is a complex notion of how a particular document relates to a given line of inquiry." Scott F. Burson, *A Reconstruction of* Thamus—*Comments on the Evaluation of Legal Information Retrieval Systems,* 79 Law Libr. J. 133, 141 (1987).

and (4) each of these systems is based on indexing and classifying that has been done by human indexers.[25]

¶12 In *The Curse of Thamus*, Dan Dabney reviewed the limitations of pre-coordinated indexing systems such as the West Digest System.[26] There is an inherent limitation on the depth of the indexing (every idea in the case cannot be indexed) and even in the complex five-tiered level of indexing in the West Digest System, the headnote exists in only one or at most two places:

> This short review of ideas in indexing shows that the indexing process is prone to many errors and uncertainties. Manual indexing is only as good as the ability of the indexer to anticipate questions to which the indexed document might be found relevant. It is limited by the quality of its thesaurus. It is necessarily precoordinated and thus also limited in its depth. Finally, like any human enterprise, it is not always done as well as it might be.[27]

¶13 The perils of relying on computer-generated searches in very large databases are equally well-documented, and finding *all* the relevant documents is a major problem.[28] Very few studies have been able to determine recall (the number of relevant documents retrieved compared to the number of relevant documents in the database),[29] because it is usually impossible to determine the total number of relevant documents in huge databases. In the few studies of size-limited databases, where the number of potentially relevant documents was known, such as Dabney's study and the STAIRS study, recall has been very poor: in both studies, recall was about twenty percent.[30] The difficulties of creating effective searches and the literalness of search engines are also problems identified with using full-text databases.[31]

---

25. Stephen M. Marx, *Citation Networks in the Law,* 10 JURIMETRICS J. 121, 122–23 (1970). Marx thought computer-generated systems based on key words alone would suffer from similar defects and proposed a context-based and citation-based retrieval system he characterized as a form of "exhaustive shepardization" assisted by a technique for "automatically isolating the factual content of a case" *Id.* at 125, 137. The article was written, of course, before Shepard's went online or KeyCite was launched.

26. Daniel P. Dabney, *The Curse of Thamus: An Analysis of Full-Text Legal Document Retrieval*, 78 LAW LIBR. J. 5, 13–14 (1986).

27. *Id.* at 14.

28. Whether finding *all* of the relevant documents is necessarily the holy grail of online research is an entirely different question. *See, e.g.,* Burson, *supra* note 24, at 136–39. In fact, whether a researcher wants high recall or high precision may vary from search to search: for comprehensive research, a searcher may need every relevant document, while for a time-driven request, the user may want the system to return the documents that are most highly relevant first. Anton Geist, Using Citation Analysis Techniques for Computer-Assisted Legal Research in Continental Jurisdictions 19 (May 1, 2009) (unpublished LL.M. thesis, University of Edinburgh), *available at* http://ssrn.com/abstract=1397674. One problem for inexperienced researchers seems to be a belief in the high quality of their research results. Novice searchers believe they have actually seen all the relevant documents and that the documents seen are the most relevant documents. Any legal research teacher can confirm this phenomenon.

29. Paul Callister's definitions of recall and precision are simple and understandable: "Essentially, there are two conflicting standards for measuring the success of your research. Precision measures how many documents were on point within your search results. In contrast, recall gauges the relevant documents in your results compared to what you could have found." Paul D. Callister, *Working the Problem,* ILL. B.J., Jan. 2003, at 43, 44.

30. Geist, *supra* note 28, at 15, 29.

31. *Id.* at 15.

¶14 When Dabney published his findings, the objections that were made at the time by representatives of both LexisNexis and Westlaw were that discussing recall and precision in the abstract failed to take into account the many value-added features of both databases.[32] The question is further complicated by the fact that defining what is *relevant* affects the results of the study: the Dabney and Gerson[33] studies used a "rigorous laboratory approach, i.e., cases are judged relevant only if they meet those on a pre-defined list . . . ."[34] Dean Mason's more recent study used a subjective standard for relevance, based on actual attorney inquiries.[35]

¶15 Precision is the number of returned documents in a search that are judged relevant. Precision is much easier to measure, once the standard of relevance has been chosen. In the large-scale studies that have been performed in legal databases, precision has varied. The results of these studies were summarized by Mason and reprinted here in slightly different form as table 1.[36]

Table 1

Review of Precision Studies

Study	Precision	
STAIRS	79%	
Dabney	12.4% Westlaw	11.5% LexisNexis
Gerson	37% Westlaw	31% LexisNexis
Mason	81% Westlaw	74% LexisNexis

¶16 Of course, the higher the precision, the poorer the recall: meaning that you will miss more relevant documents. The inverse relationship between precision and recall is "a universal principle of information science."[37] And when searching in a database with millions of documents, even high precision still results in unacceptable levels of irrelevant documents to review. The problem of poor recall can be a major hurdle, as certain types of legal research require finding all potentially rele-

---

32. Jo McDermott, *Another Analysis of Full-Text Legal Document Retrieval,* 78 Law Libr. J. 337, 343 (1986); Craig E. Runde & William H. Lindberg, The Curse of Thamus*: A Response,* 78 Law Libr. J. 345, 345–46 (1986).

33. Kevin Gerson, *Evaluating Legal Information Retrieval Systems: How Do the Ranked-Retrieval Systems of WESTLAW and LEXIS Measure Up?,* Legal Reference Services Q., 1999, no. 4, at 53.

34. Dean Mason, *Legal Information Retrieval Study—Lexis Professional and Westlaw UK*, 6 Legal Info. Mgmt. 246, 248 (2006); *see also* Sutton*, supra* note 24, at 187 (discussing the concept of relevance in actual case law retrieval: "a relevant case is one that plays some cognitive role in the structuring of a legal argument or the framing of legal advice.").

35. Mason*, supra* note 34, at 248. Mason used research requests from lawyers as queries for his study, and examined the first ten results of fifty separate searches. *Id.* at 247.

36. *Id.* at 248 tbl.2. *See also* Burson*, supra* note 24, at 135 (discussing factors affecting precision and recall in legal information retrieval systems).

37. Callister, *supra* note 29, at 44 (2003). As far back as 1994, West's own study of the relationship between precision and recall in the *Federal Supplement* database showed that as precision went up, recall went down at almost the identical rate. *Id.* at 44 & fig.2.

vant cases; failure to find an important case, either supporting or undercutting your client's position, would be a significant flaw.

The Relevance of Results Using Classification Systems

¶17 Once you have good seed cases, the online use of digests and citators should be effective ways to generate a large group of potentially relevant application cases for the researcher. I decided to first test the classification systems on LexisNexis and Westlaw to see what the results were in terms of relevance. Using both a human-generated topic system and an algorithmically-generated topic system to find "relevant" cases, one would expect some cases returned in the search sets to be different, but would hope for substantial overlap. And in the best of all possible worlds, for the unique cases in each system, each type of search would generate relevant cases not located by the other method of searching. The differences in results should also illustrate some of the benefits and detriments of each kind of searching and support the idea that to do exhaustive research, researchers may need to take advantage of both human and computer-generated indexing.

¶18 I first tested cases using the West key numbers and the Lexis Topics or "More Like This Headnote"[38] functions for similar pairs of headnotes, in order to find more cases on that particular legal topic. I then tested the use of headnotes as limiters in KeyCite and Shepard's to find more cases that cited my case for the issue in that particular headnote. Although the headnotes in Westlaw and Lexis are generated differently, in each case I reviewed there were one or more pairs of headnotes that were similar enough to make such a comparison possible.[39]

¶19 I started with a set of cases that "made new law." The cases I chose are listed below, and the numbers of the headnotes I compared are shown in table 2. To illustrate the process, the headnotes I compared from *New York Times v. Sullivan* are shown in figures 3 and 4.

1. *New York Times v. Sullivan*, 376 U.S. 254 (1964) (in a suit for defamation or libel involving public officials or public figures, plaintiff must establish that the publisher acted with actual malice or in reckless disregard of a statement's truth or falsity)
2. *Brown v. Board of Education*, 349 U.S. 294 (1955) (overturning the constitutionality of separate but equal public educational facilities and holding that *de jure* racial segregation is a violation of the Equal Protection Clause)
3. *Buckley v. Valeo*, 424 U.S. 1 (1976) (interpreting a federal campaign finance law, the Court upheld statutory limits on individual contributions, but struck down those sections of the statute that limited campaign expenditures, which were, as money spent to influence elections, a constitutionally protected form of speech)
4. *Escola v. Coca-Cola Bottling Co.*, 150 P.2d 436 (Cal. 1944) (to establish *res ipsa*

---

38. "More Like This Headnote" is discussed in detail *infra* ¶ 29.

39. The relevance checking for this article was performed between Mar. 8, 2009 and May 8, 2009 for the first five citations; the rest of the cases were checked between July 5, 2009 and Aug. 24, 2009.

loquitur, it is not necessary that the instrument causing the injury be under the exclusive control of the defendant at the time of the accident)

5. *Terry v. Ohio*, 392 U.S. 1 (1968) (upholding the constitutionality of "stop and frisk," where a police officer stops a suspect on the street and does a quick search for weapons, even without probable cause for arrest, if the police officer has a reasonable suspicion that the person has committed, is committing, or is about to commit a crime)

6. *Tarasoff v. Regents of the University of California*, 551 P.2d 334 (Cal. 1976) (if a psychotherapist has knowledge that a patient is threatening harm to a third party, the psychotherapist has a duty to protect that person, either by notifying police, warning the intended victim, or taking other reasonable steps to protect the threatened individual)

7. *San Diego Federal Credit Union v. Cumis Insurance Society, Inc.,* 208 Cal. Rptr. 394 (Cal. Ct. App. 1984) (where an insurer has reserved its rights to deny coverage under an insurance policy, the insurer must explain the implications of the insurer representing its own interests and the interests of the insured, and provide the insured with independent counsel, at the insurer's expense)

8. *Foley v. Interactive Data Corp.*, 765 P.2d 373 (Cal. 1988) (at-will employment can be altered by actions, such as regular promotions, verbal assurances of job security, salary increases, and employee termination guidelines, resulting in an implied-in-fact contract not to terminate employment except for cause)

9. *Golden Gateway Center v. Golden Gateway Tenants' Ass'n.*, 29 P.3d 797 (Cal. 2001) (the ban by a private owner of an apartment complex against distribution of a tenant association newsletter involved no state action and thus did not violate the First Amendment)

¶20 For each pair of headnotes, I created a statement of relevance that would guide my review of each case (shown in table 2). I chose my criteria for relevance prior to reviewing any cases. Each statement of relevance is taken directly from the headnote, but is slightly more factually focused, so it would more closely parallel a real-world research need. To review cases for relevance, I read the summary of each case and, where necessary, used Focus (LexisNexis) or Locate in Result (Westlaw) to review the relevant portions of the opinion before making a determination.[40] For example, for the chosen headnotes of *New York Times v. Sullivan*, the standard of relevance was whether the case discussed factual circumstances in which police officers are "sharply attacked." There are two topic lines for the LexisNexis headnote, and I chose the second line because it was more focused on my relevance criteria—a general overview on commercial speech was not what I wanted.

---

40. As always, determinations of relevance are subjective. As Scott Burson has noted, "we can profoundly disagree on the relevance of a particular document." Burson, *supra* note 24, at 141. I reviewed these cases as I did when I was an attorney, looking for factors that would make me put it in the "this looks like it might be useful" file.

West Key Numbers

¶21 There were 107 headnotes associated with West key number 92k1625 when I searched state and federal cases. To make sure that I only reviewed the headnotes with this key number that were about other situations in the federal courts in which police officers had been "sharply attacked," I limited the jurisdiction to federal cases and added to my search the terms (**(police law peace) /2 officer!**).[41] The search screen for the Custom Digest can be seen in figure 5.

Table 2

Cases Used to Compare West Topics and Key Numbers, Lexis Topics,
and "More Like This Headnote"

Case Citation and Headnote Numbers	Criteria for Relevance
1. *New York Times v. Sullivan*, 376 U.S. 254 (1965) – LN 9 and WL 9	Discusses circumstances in which police officers are "sharply attacked."
2. *New York Times v. Sullivan*, 376 U.S. 254 (1965)– LN 4 and WL 4	Discusses constitutional protection of speech that had a commercial or advertising aspect regardless of depth of treatment.
3. *Brown v. Bd. of Education*, 349 U.S. 294 (1955) – LN 7 and WL 8	Discusses the implementation of desegregation plans and the interplay between constitutional and administrative requirements.
4. *Buckley v. Vallejo*, 424 U.S. 1 (1976) – LN 6 and WL 16, 17	Discusses limitations on the right of association in the election campaign setting.
5. *Escola v. Coca-Cola Bottling Co.*, 150 P.2d 436 (Cal. 1944) – LN 3 and WL 2	Discusses the factual circumstances where an article that has left the defendant's control is changed or unchanged (element of *res ipsa loquitur*).
6. *Terry v. Ohio*, 392 U.S. 1 (1968) – LN 15 and WL 23	Discusses when officer was or was not justified in searching for weapons (in the absence of a warrant).
7. *Tarasoff v. Regents of the Univ. of Calif.*, 551 P.2d 334 (Cal. 1976) – LN 15 and WL 35	Discusses the parameters of those covered by statute regarding detainment of persons for 72 hours and concomitant lack of liability.
8. *San Diego Fed'l v. Cumis Ins. Soc'y*, 208 Cal. Rptr. 494 (Cal. Ct. App. 1984) – LN 7 and WL 4)	Discusses the types or scope of coverage issues that raise a conflict requiring appointment of *Cumis* counsel.
9. *Foley v. Interactive Data Corp.*, 765 P.2d 373 (Cal. 1988) – LN 13 and W L4	Discusses the variations of factual circumstances that will or will not take an employment agreement out of the "at-will" category into the "for cause" category.
10. Golden Gateway Center v. Golden Gateway Tenants' Ass'n, 29 P.3d 797 (Cal. 2001) – LN 26 and WL 3	Discusses when a place is sufficiently open (or closed) to the public to implicate state action.

LN=LexisNexis; WL=Westlaw

---

41. Because I was comparing results from headnotes (West Key Numbers) and full-text results (Lexis Topics), I tried to choose keyword searches that would not be too limiting when only searching headnotes and would also not be too broad when searching the full text of cases on LexisNexis (it is not possible to search only headnotes on LexisNexis). When it was unclear from the headnote whether or not the case was relevant, I would review the full text of the case, and use the summary of the case and, if necessary, "Locate in Result," to determine relevance.
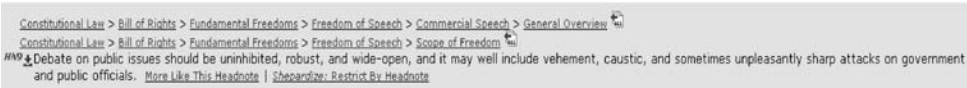
Figure 3. Headnote from *New York Times v. Sullivan* on LexisNexis
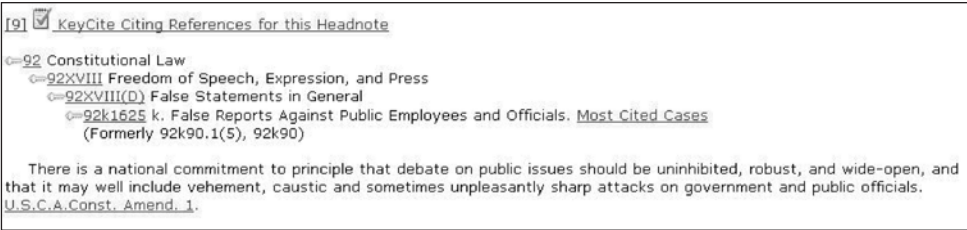
Figure 4. Headnote from *New York Times v. Sullivan* on Westlaw
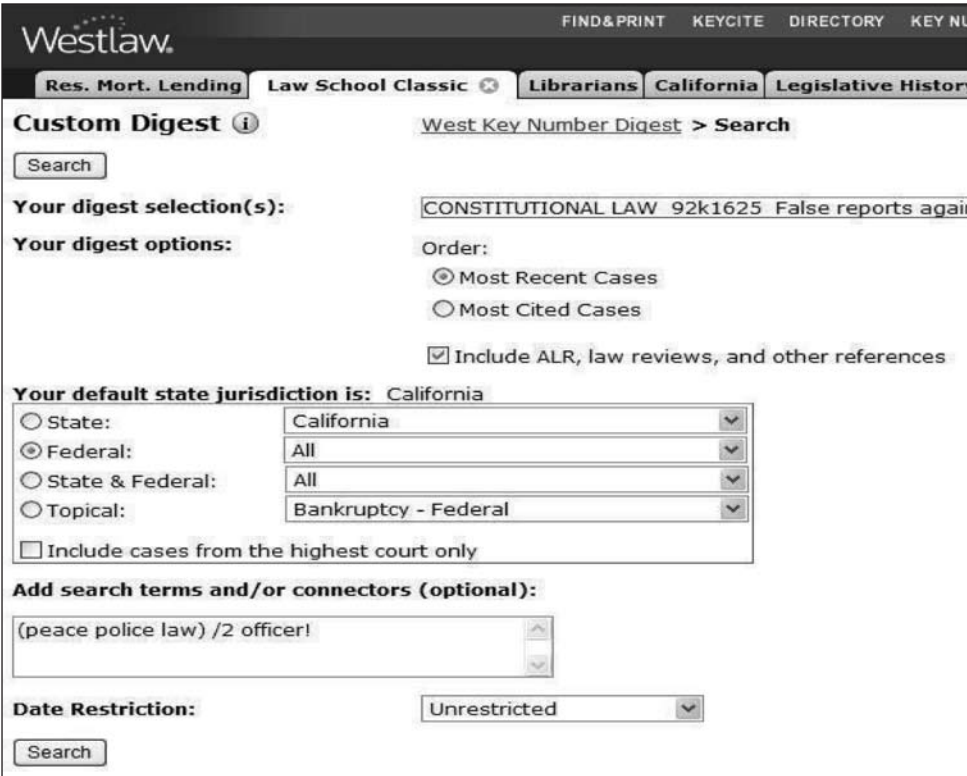
Figure 5. Custom Digest Search Screen on Westlaw

There were six results, and five unique citations.[42] Relevance was determined by a discussion of a situation in which police have been sharply attacked. All of the cases were relevant:

1. *Chaker v. Crogan*, 428 F.3d 1215 (9th Cir. 2005)
2. *Hamilton v. City of San Bernardino*, 325 F. Supp. 2d 1087 (C.D. Cal. 2004)
3. *Eakins v. Nevada*, 219 F. Supp. 2d 1113 (D. Nev. 2002)
4. *Hamilton v. City of San Bernardino*, 107 F. Supp. 2d 1239 (C.D. Cal. 2000)
5. *Gates v. City of Dallas*, 729 F.2d 343 (5th Cir. 1984)
6. *Gates v. City of Dallas*, 729 F.2d 343 (5th Cir. 1984)

Lexis Topics

¶22 For *New York Times v. Sullivan*, in LexisNexis the most relevant Lexis Topic path is: *Constitutional Law > Bill of Rights > Fundamental Freedoms > Freedom of Speech > Scope of Freedom*. There are a number of options from this screen, which can be seen in figure 6. One is to click the "all" button next to the Lexis Topic path, which the mouse-over indicates "retrieve[s] all headnotes and additional cases on this topic."
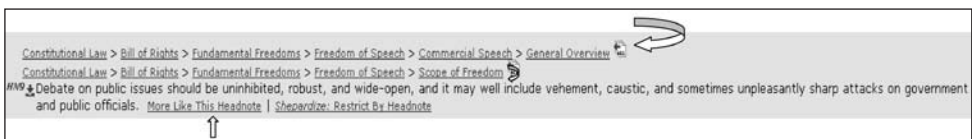
Figure 6. "All" Icon Used for Searching Topics on LexisNexis

If you click on the "all" button, you must then select a jurisdiction.[43] I selected "Federal Constitutional Law Cases" and after searching received an error message: "More than 3000 cases." So even though the case results were limited to constitutional law cases, there were more than 3000 results for this Lexis Topic generally (compared to 107 from the West key number).

¶23 A searcher can also click on the lowest link in the Lexis Topic chain—"Scope of Freedom," which brings up the screen shown in figure 7. Clicking on the "Scope of Freedom" link on this page gives a searcher two options, shown in figure 8.

---

42. In the West digest results, there may be multiple entries for one case, if there is more than one headnote in the case with the same key number. Where multiple entries were retrieved, each entry was counted separately.

43. See *infra* figure 8.

Search | Research Tasks | Get a Document | *Shepard's®* | Alerts | Total Litigator | Transactional Adv

by Source | **by Topic or Headnote** | by Guided Search Form | by Dot Command

All Topics

Look for a Legal Topic

Option 1: Find a Legal Topic

[   ] 🔍 Find

Find in: ⦿ Constitutional Law
     ○ All Topics

Option 2: Explore Legal Topics
⊟ ⓘ **Constitutional Law**
   ⓘ **General Overview**
   ⊕ ⓘ **Congressional Duties & Powers** (Related Topics...)
   ⊕ ⓘ **The Presidency**
   ⊕ ⓘ **The Judiciary** (Related Topics...)
   ⊕ ⓘ **Relations Among Governments**
   ⓘ **Separation of Powers** (Related Topics...)
   ⓘ **Amendment Process**
   ⊕ ⓘ **Supremacy Clause** (Related Topics...)
   ⊖ ⓘ **Bill of Rights**
     ⓘ General Overview
     ⊖ ⓘ Fundamental Freedoms (Related Topics...)
       ⓘ General Overview
       ⓘ Freedom of Assembly
       ⓘ Freedom of Association (Related Topics...)
       ⊕ ⓘ Freedom of Religion (Related Topics...)
       ⊖ ⓘ Freedom of Speech (Related Topics...)
         ⓘ General Overview
         ⓘ Advocacy of Illegal Action
         ⊕ ⓘ Commercial Speech
         ⊕ ⓘ Defamation
         ⓘ Expressive Conduct
         ⓘ Fighting Words
         ⓘ Forums
         ⊕ ⓘ Free Press (Related Topics...)
         ⓘ Obscenity (Related Topics...)
         ⓘ Political Speech (Related Topics...)
         ⓘ Public Employees (Related Topics...)
         ⓘ Scope of Freedom

Figure 7. Screen to Select a Topic on LexisNexis

All Topics > Constitutional Law > Bill of Rights > Fundamental Freedoms > Freedom of Speech > **Scope of Freedom** ⓘ (Browse topic index)

Option 1 - Search across Sources ?

Search for 'Scope of Freedom' across Cases, Statutes, Analysis and more...

① Jurisdiction
[Federal ▾] Select Multiple...

② Sources
[Select Sources...]

③ Search Terms (Optional)
⦿ Terms and Connectors ○ Natural Language

[376 us 254] 🔍 Search
Suggest Terms
for My Search

Search Connectors   **How Do I...?**
and   and   › Restrict by date on the Search by Topic form?
or   or   › Search across different sources?
w/N   within N words
pre /N   precedes by N words
w/p   in same paragraph
›More Connectors & Commands...   🗎 **View Topic Search Tutorials**

OR

Option 2 - Search by Headnote ?

Retrieve all headnotes and additional cases on 'Scope of Freedom'

Jurisdiction
[Select a Jurisdiction ▾]

Date (Optional)

⦿ No Date Restrictions ▾

○ From [   ] To [   ] Date Formats...

🔍 **Retrieve All**

What's this?

Retrieve all headnotes classified to the topic or cases that discuss the topic, even if there isn't a specific headnote.

Figure 8. Screen Shown on LexisNexis When a Topic is Chosen from Topic Outline

¶24 Option 2, "Search by Headnote," duplicates the search that can be performed by clicking the "all" button and limiting the results to "Federal Constitutional Law Cases." This search also results in retrieval of an error message: "More than 3000 results." Because there is no way to further limit the results prior to searching, the search cannot be run using this option. When I accidentally ran this search limiting the results to California, I got 1719 results, and the option to "Show Headnotes Only," which the mouse-over tag told me will "limit list to cases with headnote on the topic." The "Show Headnotes Only" option confirms that the results set from this search includes cases that do *not* reference the relevant headnote.[44]

¶25 The box on the left, "Option 1: Search Across Sources," appeared with the case citation already printed in the box. I added the search string: (**(police or peace or law**) **/2 officer!**). Although this would appear to limit the results to cases that cite your case, I allowed it to remain, because without the citation limiter I was unable to run the search. I received a "More than 3000 results" warning again. From the drop-down menu I selected "Federal" as the jurisdiction and from the sources menu, I chose "Federal Constitutional Law Cases." With this search I was able to narrow the results to 241. For purposes of my review (and for most researchers' purposes), this was too many cases to review. However, hoping that the results were returned in order of relevance,[45] and to limit my result set to a number roughly comparable to the West results, I reviewed the first ten cases.[46] The results page is shown in figure 9.

¶26 None of the ten cases (listed below) were relevant to my need—cases that discuss a situation in which police have been "sharply attacked."

1. *Illinois ex rel. Madigan v. Telemarketing Assocs.*, 538 U.S. 600 (2003)
2. *Bartnicki v. Vopper*, 532 U.S. 514 (2001)
3. *Denver Area Educ. Telcoms. Consortium v. FCC*, 518 U.S. 727 (1996)
4. *R. A. V. v. St. Paul*, 505 U.S. 377 (1992)
5. *Florida Star v. B. J. F.*, 491 U.S. 524 (1989)

---

44. When I limited my results of my accidental search by "Show Headnotes Only," the set of 678 results at least contained my headnote. At this point I could have used the Focus bar, or clicked on an "In Depth Discussion" link. When I chose Focus, my results were reduced to 292 cases, but the "In Depth Discussion" link was no longer available. The first ten results were not relevant to my research. The second choice was to first choose the "In Depth Discussion" link, which the mouse-over tag informs me will "retrieve up to 15 discussion cases on this topic." With those fifteen cases, when I added my search string ((**police or law or peace**) **/2 officer!**), I retrieved two cases, neither of which was relevant.

45. It appears from the results that LexisNexis automation ranks these result sets first according to the occurrence of terms matching the limiting search string that occur *anywhere* in the text of retrieved cases, and *then* according to headnotes that match the topic. *See infra* ¶ 28 (discussing the location of relevant cases found using the West key number search in LexisNexis results sets). Thus, results are not limited to cases that include the relevant headnotes.

46. For LexisNexis results sets that returned hundreds of documents, if the Westlaw results set was less than ten, I limited my review to the first ten results; for Westlaw results sets of more than ten, I reviewed the same number of LexisNexis results. The limitations of time prevented me from reviewing the entire results sets. A researcher with a real-world problem would have to determine the number of results to review based on client need.
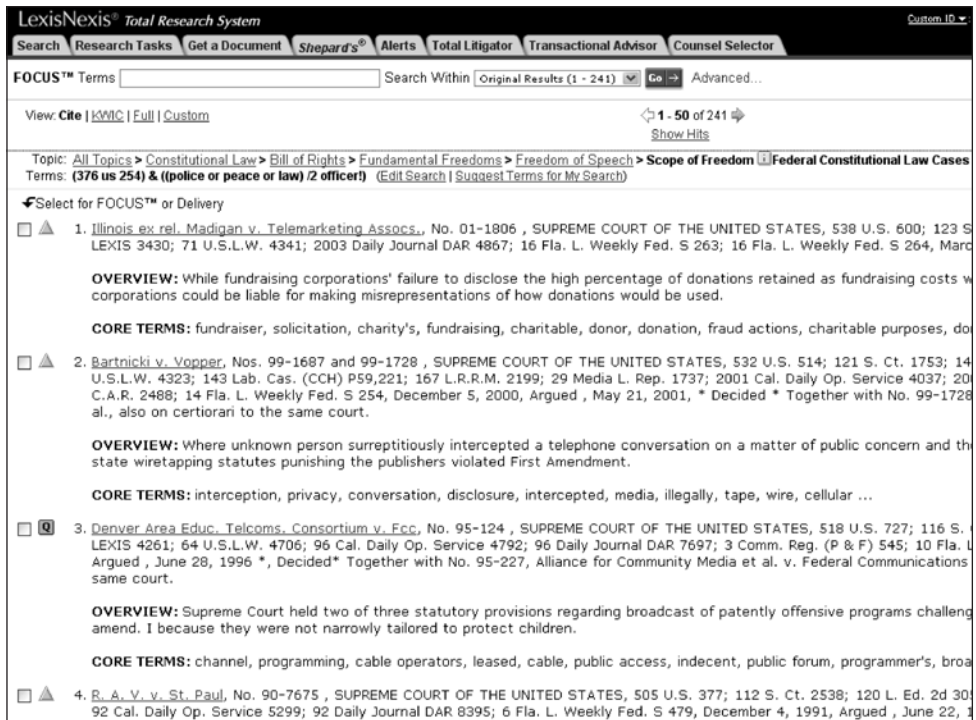
Figure 9. Search Results Screen on LexisNexis

 6. *Riley v. Nat'l Federation of the Blind*, 487 U.S. 781 (1988)
 7. *Frisby v. Schultz*, 487 U.S. 474 (1988)
 8. *Rankin v. McPherson*, 483 U.S. 378 (1987)
 9. *Members of City Council v. Taxpayers for Vincent*, 466 U.S. 789 (1984)
 10. *NAACP v. Claiborne Hardware Co.*, 458 U.S. 886 (1982)

¶27 I then decided to compare the two sets of results by looking at the list of 241 Lexis Topic results for the five relevant cases I found using the West key number 92k1625. Their locations in that list are listed below:

 1. *Chaker v. Crogan*, 428 F.3d 1215 (9th Cir. 2005); this case was 36/241.
 2. *Hamilton v. City of San Bernardino*, 325 F. Supp. 2d 1087 (C.D. Cal. 2004); this case was not part of the LexisNexis result set.
 3. *Eakins v. Nevada*, 219 F.Supp.2d 1113 (D. Nev. 2002); this case was 148/241.
 4. *Hamilton v. City of San Bernardino*, 107 F. Supp. 2d 1239 (C.D. Cal. 2000); this case was 107/241.
 5. *Gates v. City of Dallas*, 729 F.2d 343 (5th Cir. 1984); this case was 98/241.

¶28 The fact that the relevant cases found on Westlaw were, if they appeared at all, scattered throughout the LexisNexis results, among many irrelevant results,

indicates that the language of the LexisNexis headnote is *not* the main determinant of relevance in the LexisNexis result set. If the language were the determining factor of relevance, you would expect relevant results to all be grouped at the beginning of the LexisNexis results. What is puzzling is why the default ordering for the LexisNexis result would be to list first results that do not include the language of the relevant headnote, as these are unlikely to be relevant to a searcher who chooses this pathway.

"More Like This Headnote"

¶29 There is a third option to search in LexisNexis—"More Like This Headnote." LexisNexis Research Help describes this option as follows: "Click the 'More like this Headnote' link at the end of a LexisNexis headnote to see all the cases with LexisNexis Headnotes related to that specific headnote." [47] I took this to mean that it would find all cases related to the headnote topic whether the cases cite *New York Times v. Sullivan* or not. Using this option took me to the screen shown in figure 10, and, when limited to Combined Federal Courts, returned 250 results, as shown in figure 11. [48] The first 250 results were sorted by "Closest Match." After running the search, the searcher can use Focus to search within the results. I focused on (**peace or police or law**) **/2 officer!**). [49] There were sixty-four results, and the results screen is shown in figure 12.

¶30 The first ten cases from the results are listed below. The starred cases were *not* relevant; there were three relevant results in the first ten: [50]

1. *United States v. Carrier*, 517 F. Supp. 644 (N.D.N.Y. 1981)
2. *United States. v. Olson*, 629 F. Supp. 889 (W.D. Mich. 1986)
3. *McKinley v. Baden*, 777 F.2d 1017 (5th Cir. 1985)
4. *Fogel v. Collins*, 531 F.3d 824 (9th Cir. 2008)
5. *Jones v. Key West*, 679 F. Supp. 1547 (S.D. Fla. 1988)
6. *Johnson v. Robbinsdale Independent School District No. 281*, 827 F. Supp. 1439 (D. Minn. 1993)
7. *United States v. Poocha*, 259 F.3d 1077 (9th Cir. 2001)
8. *Flynn v. Giarrusso*, 321 F. Supp. 1295 (E.D. La. 1971)
9. *Smith v. Twp. of Aleppo*, 2005 U.S. Dist. LEXIS 44219 (W.D. Pa. July 13, 2005)
10. *Ferris v. Larry Flynt Publishing, Inc.*, 2001 U.S. Dist. LEXIS 7759 (D. Haw. Apr. 13, 2001)

---

47. LexisNexis, Research Help: More Like This, http://web.lexis.com/help/research/gh_browsing .asp#MoreLikeThis (last visited Feb. 14, 2010).

48. I have my preferences set in LexisNexis as follows: the default for natural language results in LexisNexis is set to 250 results and the results are returned in order of relevance.

49. This search was run originally in May 2009 and rerun on July 5, 2009, and there were a few differences in the results. The second original results set was 250, rather than 249. Also, the "sort by closest match" button was new; that button had not been there in May. "Sort by closest match" confirms the algorithmic nature of the search and the ordering of the results; it does not explain why the results are so different from the Lexis Topic search. After searching using my Focus terms, I reviewed all sixty-four cases I found in July, and compared them to the ten I reviewed in May; of the 64, 9 were relevant (14%). The additional relevant cases were numbers 18, 21, 35, 39, 45, and 46 in the list ordered by "closest match."

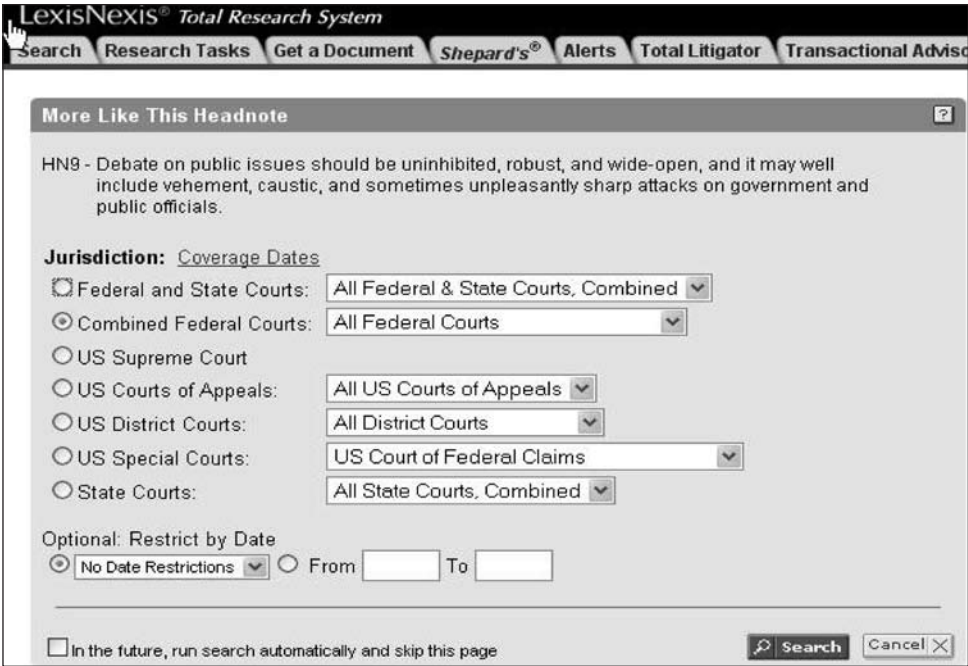50. Relevance was again determined by whether cases discussed a situation in which police have been "sharply attacked."

Figure 10. "More Like This Headnote" Search Screen on LexisNexis
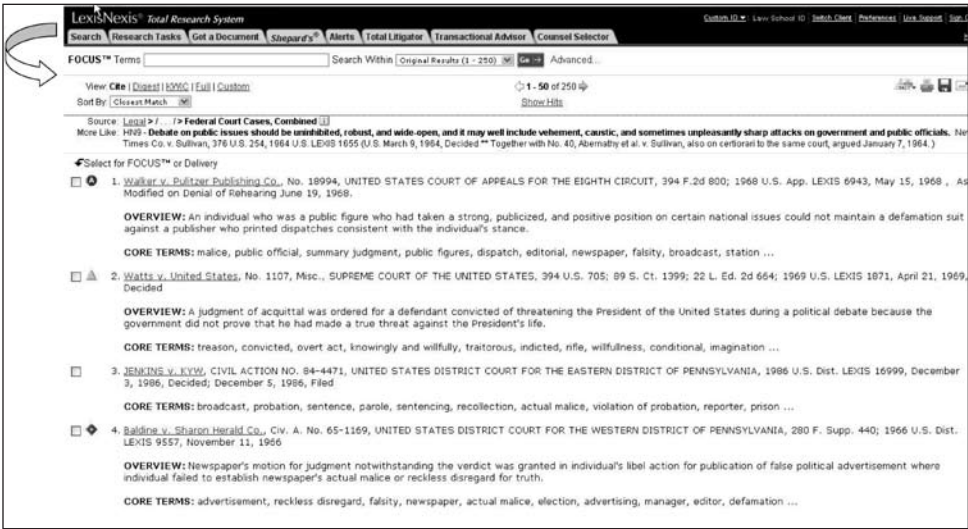
Figure 11. Sample Results Screen from "More Like This Headnote" Search

Figure 12. Focus Used to Limit "More Like This Headnote" Search

¶31 What is significant about these results is that the first ten cases are completely different from the first ten cases listed when I searched the "Scope of Freedom" topic. I thought that perhaps the difference might be because I was using the second classification line from Lexis Topics (*Constitutional Law > Bill of Rights > Fundamental Freedoms > Freedom of Speech > Scope of Freedom*), and "More Like This Headnote" was matching cases from the first classification line in Lexis Topics (*Constitutional Law > Bill of Rights > Fundamental Freedoms > Freedom of Speech > Commercial Speech > General Overview*). I ran the search again using the "*General Overview*" link from the first classification line. In "Option 1—Search Across Sources," I limited my jurisdiction to "Federal," chose "Federal Constitutional Law Cases" as my source, and used the same terms and connectors Focus search. There were fifty-three results—not sixty-five as there were using the "More Like This Headnote" search. The first ten cases were completely different from the ten cases from my "More Like This Headnote" search. Each of these searches generated different results, even though the mechanics of topic assignment would indicate that the headnote and the topic should cover similar territory. There simply did not appear to be a strong correlation between the LexisNexis headnote and the Lexis Topic.

¶32 Though complex, what all these comparisons come down to is that there are *at least* three potential results sets: one for West key numbers, one for Lexis Topics, and one for LexisNexis's "More Like this Headnote." Because Lexis headnotes frequently have more than one topic entry, each entry line is another potential set of results, and each of these results sets is different.

¶33 I also checked to see if the five relevant cases I found using West key number 92k1625 were in the sixty-five LexisNexis "More Like This Headnote" results:

1. *Chaker v. Crogan*, 428 F.3d 1215 (9th Cir. 2005); this case was 44/65.
2. *Hamilton v. City of San Bernardino*, 325 F. Supp. 2d 1087 (C.D. Cal. 2004); this case was not part of the LexisNexis result set.
3. *Eakins v. Nevada*, 219 F. Supp. 2d 1113 (D. Nev. 2002); this case was 28/65.
4. *Hamilton v. City of San Bernardino*, 107 F.Supp.2d 1239 (C.D. Cal. 2000); this case was not part of the LexisNexis result set.
5. *Gates v. City of Dallas*, 729 F.2d 343 (5th Cir. 1984); this case was not part of the LexisNexis result set.

¶34 I performed similar comparisons for nine other pairs of headnotes. The results for all the cases I reviewed are set forth in table 3.[51] Any relevant information on search limits is noted in table 3.1. The average percentage of relevant cases found using West key numbers, "More Like this Headnote," and Lexis Topics is shown in figure 13.
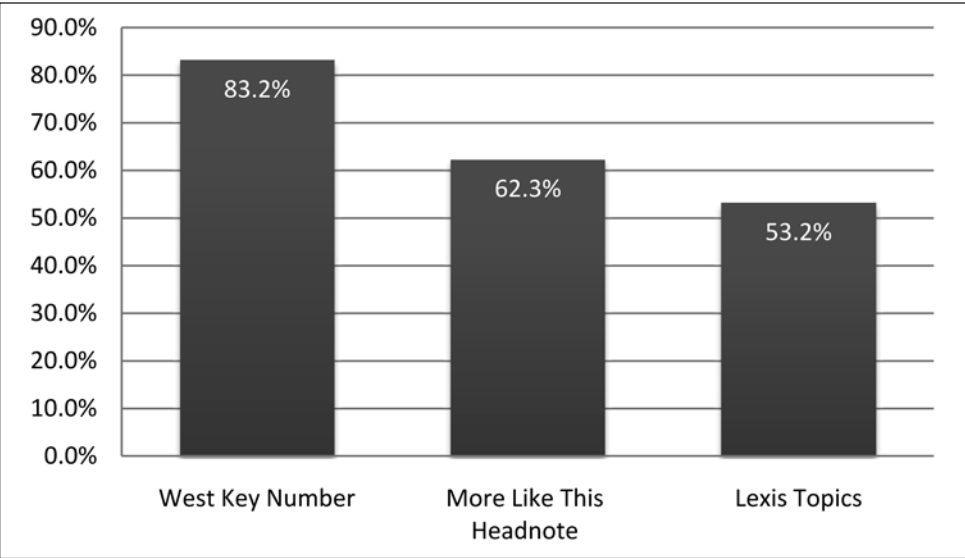
Figure 13. Average Percentage of Cases Found That Were Relevant

Using Classification Systems for Research

¶35 This study was of necessity small, since each headnote comparison was very time consuming. The percentage of relevant cases might very well change if a larger pool of matching headnotes were compared, and I invite others to continue enlarg-

---

51. To make the results as comparable as possible, and because a default of 250 was simply too many cases to review, where there were up to ten Westlaw results, I reviewed the first ten results on "More Like This Headnote" and the Lexis Topic; if there were more than ten results on Westlaw, I reviewed the same number of cases in "More Like This Headnote" and the Lexis Topic.

ing the study. But even leaving absolute percentages aside, the comparisons in table 3 reveal some strong trends. One thing you can clearly say about using the West and LexisNexis features is that each one will find some relevant cases on your legal topic. However, it does not look like any one of them can be used to find all relevant cases on your legal topic.

¶36 Judging only from the cases I reviewed, West's key numbers, on average, appear to deliver more relevant results than either of the LexisNexis options. Because a major difference between the two systems seems to be the degree of dependence on algorithms for creating classification topics and assigning head-notes to each topic,[52] the role of human editors appears to be a definite advantage in returning relevant results. It seems that it is only the Custom Digest results that are limited to those cases where the language of the headnote is present, and if limiting terms are added to the classification search, where those limiting terms also appear. Even with this advantage, the LexisNexis classification scheme still returns relevant cases not found using the key number system. The percentages of *unique* relevant cases found using each type of search is shown in figure 14.

¶37 After reviewing the LexisNexis results sets, including the way the results are displayed, the choices for display of results, and the very different results for all of the LexisNexis result sets, I was convinced that the LexisNexis results are generated
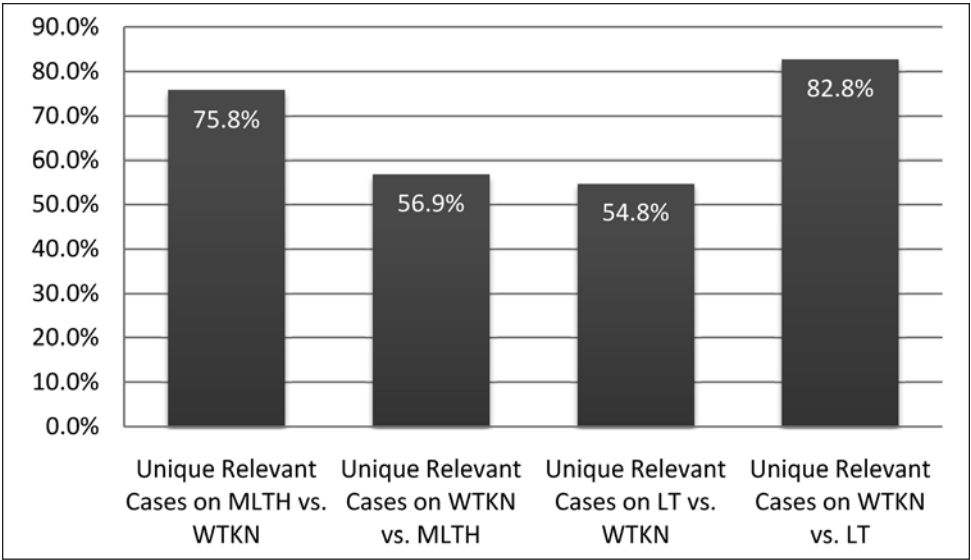
Figure 14. Average Percentage of Relevant Cases Found
That Were Not Found Using Other System

WTKN=West Topic & Key Number; MLTH=LexisNexis "More Like This Headnote"; LT=Lexis Topic

---

52. Of course, the headnotes themselves are created differently, with West editors reframing the legal issues in each case in their own language and LexisNexis taking the headnotes directly from the language of the case (*see supra* ¶¶ 5–6), but because the headnotes being compared for this test were chosen for their similarity, the method of generation should not really matter.

Table 3

Relevant Cases Found with West Topics & Key Numbers, Lexis Topics, and "More Like This Headnote"

Citation	WTKN # Cases Retrieved (Topic & Key Number Used)	% Cases Relevant	MLTH # Cases Retrieved/Reviewed (Headnote number used)	% Relevant Cases	# (%) Unique[a] Relevant Cases on MLTH vs. WTKN	# (%) Relevant Cases on WTKN vs. MLTH	LT # Cases Retrieved/Reviewed (Lexis Topic)	% Relevant Cases	# (%) Unique Relevant Cases on LT vs. WTKN	# (%) Unique Relevant Cases on WTKN vs. LT
1. 376 U.S. 254	6 (92k1625)	100% (6/6)	65/10 (HN 9)	30% (3/10)	3 (100%)	3 (50%)	241/10 (Scope of Freedom)	0% (0/10)	0 (0%)	6 (100%)
2. 376 U.S. 254	1 (92k2164)	100% (1/1)	170[b]/10 (HN 4)	10% (1/10)	1 (100%)	0 (0%)	133/10[c] (Commercial Speech–Overview)	10% (1/10)	1 (100%)	1 (100%)
3. 349 U.S. 294	8 (345k13(10))[d]	100% (8/8)	62/10[e] (HN 7)	100%[e] (10/10)	6 (60%)	4 (50%)	33/10[f] (Attendance District & Zones)	100% (10/10)	46[e] (60%)	4 (50%)
4. 424 U.S. 1	29 (92k4460)[g]	76% (22/29)	87/29 (HN 6)	52% (15/29)	12 (80%)	19 (86%)	54/29[h] (Right of Association)	17% (5/29)	4 (80%)	21 (95%)
5. 150 P.2d 436	19 (272k1614)[i]	95% (18/19)	71[j]/19 (HN 3)	84% (16/19)	13 (81%)	15 (83%)	89/19 (Overview Res Ipsa)[k]	63% (12/19)	6 (50%)	12 (67%)
6. 392 U.S. 1	25 (349k70)	96% (24/25)	70/25 (HN 15[l])	52% (13/25)	8 (61%)	19 (79%)	25/25 (Stop & Frisk Overview)	88% (22/25)	16 (73%)	18 (75%)
7. 551 P.2d 334	6 (81k7)[m]	17% (1/6)	61[n]/10 (HN 15)	80% (8/10)	7 (87%)	0 (0%)	204[o]/10 (Immunity Overview)	0% (0/10)	0 (0%)	1 (100%)
8. 208 Cal. Rptr. 494	71(217k2929)	89% (63/71)	250 /71 (HN 7)	34% (24/71)	14 (58%)	53 (84%)	20/20 (Conflict of Interest)	25% (5/20)	4 (80%)	62 (98%)
9. 765 P.2d 373	39 (231Hk40 (2))[p]	79% (31/39)	250[q]/39 (HN 13)	92% (36/39)	22 (61%)	17 (55%)	255[r] /39 (At Will Employment–Exceptions–Implied Contract)	46% (18/39)	6 (33%)	19 (61%)
10. 29 P.3d 797	35[s] (92K1876)	80% (28/35)	250/19[t] (HN 15)	89% (17/19)	12 (70%)	23 (82%)	32/32 (Freedom of Speech–Forums)	56% (18/32)	13 (72%)	23 (82%)

WTKN=West Topic & Key Number; MLTH=LexisNexis "More Like This Headnote"; LT=Lexis Topic

Table 3.1

Explanatory Information for Table 3

a. This is the number of relevant cases that were **not** found using the other system.

b. The jurisdictional limiter was federal cases. The result of clicking "More Like This Headnote" was the default 250; I used Focus to limit by: **advertis! or commerc!** and reduced my results to 170.

c. Using the jurisdiction limit "All Federal Constitutional Cases" returned a notice that said: "results are over 3000, please revise your search," so I limited my search to the 10th Circuit, despite the fact that my Westlaw search was in all federal cases. I then used Focus, and limited by: **advertis! or commerc!** to get the results down to 133.

d. The jurisdictional limit was 10th Circuit. There were three distinct cases.

e. I reviewed all 62 cases for this headnote; the full set is on file with the author. Although the first 10 cases were all relevant, the total set contained only 18 relevant cases; 44/62 (71%) were irrelevant and the cases became increasingly irrelevant as I went down the list.

f. There were eight distinct cases.

g. Only 92k1460 was run as a search. 92k1440, headnote 16's key number, is limited to strict scrutiny, and is not relevant when searched alone. Of the 20 results, there were 17 unique cases.

h. There were over 3000 references using "Retrieve all headnotes & additional cases for Freedom of Association," even limited to "All Federal Constitutional Cases" (all federal cases is not available in LexisNexis as a limiter). I then limited the search to the 10th Circuit, and got 305 results. Using Focus to limit the results to cases with: **campaign and election!**, I narrowed it to 54 results, and reviewed the first 29 for a more even comparison.

i. The jurisdictional limit was California; keyword limitation was: **control /p chang!**.

j. The jurisdictional limit was California; focus was: **control /p chang!**; 250 default results were reduced to 71.

k. The jurisdictional limit was California; focus was: **control /p chang!**; 575 results were reduced to 89.

l. The jurisdictional limit was 9th Circuit; the Focus keywords were: **weapon! /25 warrant!**. I used the same Focus search for the Lexis Topic search (jurisdiction was federal and source was 9th Circuit cases, which include district court and bankruptcy cases)

m. The jurisdictional limit was California; the keyword limiter was: **immun!**.

n. Limiting the results to California, there were 214 results. With Focus keyword: **immun!** there were 61 results.

o. Limiting the results to California, and keyword limiting with: **immun!**, there were 204 results.

p. Limited by **"at will" /p cause** and the jurisdictional limit is California.

q. Focus was: **"at will" /p cause** and there were197 results; jurisdictional limit was California.

r. Focus was: **"at will" /p cause**; the jurisdictional limit is California. Database was California State Cases. This limited the results to 255.

s. For this case, the "More Like This Headnote" search and the Lexis Topic Search both had the following limits: The jurisdiction was California, the source was California cases, and the Focus was: **"state actions" and "public forum."** For the West search of key number 92k1780 (headnote 3), the jurisdiction was California. There were no keyword limits put in the Custom Digest, but I used **"state action"** and **"public forum"** as keywords in Find in Results searches in the full text of cases where it was not clear from the annotation whether or not a case was relevant.

t. The results for "More Like This Headnote" and Lexis Topic shown are the exact number of cases returned after using Focus.

algorithmically. In the small sample I reviewed, Lexis Topics did not deliver the level of relevance that can be generated by using "More Like This Headnote." The Lexis Topic results indicate that the language of the headnote need not be present in all of the first-returned results, thus attenuating the relevance of the results that the average searcher would be willing to review. On the basis of this small sample, Lexis Topics do not appear to be an efficient or cost-effective method of finding more relevant cases. The "More Like This Headnote" function returns, on average, more relevant cases than Lexis Topics, but the results were not always returned with the most relevant results at the beginning of the result sets—as one would hope with a complex algorithmic search.

¶38 So what is the conscientious researcher to do? It is important not to treat any of the classification systems as exhaustive. Although the West Key Number System seems to have an advantage, if comprehensive research is required, either both systems must be used, or the researcher must fill in the gap by making sure that enough secondary sources have been reviewed to assure a good complement of seed cases. Using multiple seed cases, every available relevant headnote topic must to be researched—not just the single, potentially relevant headnote topic that I researched for this comparison. Each case must be Shepardized and KeyCited, as well. Hopefully, the redundant nature of legal research will eventually pull in every relevant case.

Comparison of Online Citators

¶39 My comparison of classification systems revealed large differences in results sets and differences in the average percentages of relevant unique cases returned. What happens when the results sets from citators were compared? Citators are, of course, another widely used tool to find additional cases on a legal topic. The two major competitors are Shepard's and KeyCite. Both citation systems are routinely used to check whether or not a case is still good law, but can also be used to find more cases on a legal topic. Citation checking is forward looking. Unlike classification systems, which are designed to give you cases on your legal topic regardless of the date the citing case was decided and whether or not the researcher's seed case is mentioned, citation systems look for every instance of a case that has cited the seed case. In addition to answering the question, "is my seed case still good law?", citation systems will turn up positive cases with different factual matrices that may be more relevant to the situation the researcher is investigating.

¶40 As I have discussed earlier, West and LexisNexis use different methods to generate headnotes, with West editors summarizing the legal points in a case in their own language, and LexisNexis taking the language of its headnotes directly from the language of the case. But both systems assign those headnote numbers to citing cases algorithmically. When KeyCite was created, those headnotes had to be linked to all other cases both citing the original case and citing the original case on the same headnote. This linking is done using computer algorithms.[53] When LexisNexis generates a headnote-limited Shepard's report, "the text of the headnote

---

53. McKenzie, *supra* note 14, at 85, 86–87.

of the Shepardized case is compared algorithmically with language from the citing cases to identify references (within the citing case) that match the language of the LexisNexis headnote within the Shepard's report."[54] When similar headnotes are compared, what are the results for the relevance of the reports generated by KeyCite and Shepard's?

¶41 I started with the same set of cases I used for the classification comparison—cases that "made law" and have generated a large number of citing references, allowing me to compare the headnote feature in KeyCite and Shepard's.[55] Table 4 shows the number of citing references for each of my cases.

Table 4

Total Citing Cases from Shepard's and KeyCite

Cited Case	Shepard's	KeyCite
New York Times v. Sullivan, 376 U.S. 254 (1965)	3,158	5,624
Brown v. Bd. of Education, 349 U.S. 294 (1955)	827	909
Buckley v. Vallejo, 424 U.S. 1 (1976)	2,559	2,432
Escola v. Coca-Cola Bottling Co., 150 P.2d 436 (Cal. 1944)	328	322
Terry v. Ohio, 392 U.S. 1 (1968)	30,732	27,118
Tarasoff v. Regents of the Univ. of Calif., 551 P.2d 334 (Cal. 1976)	752	730
San Diego Fed'l v. Cumis Ins. Soc'y, 208 Cal. Rptr. 494 (Cal. Ct. App. 1984)	215	202
Foley v. Interactive Data Corp., 765 P.2d 373 (Cal. 1988)	1,050	1,068
Golden Gateway Center v. Golden Gateway Tenants' Ass'n, 29 P.3d 797 (Cal. 2001)	34	37

¶42 These numbers illustrate the unmanageable number of results there can be for seminal cases on a specific topic. A reasonable researcher would want to further limit results. For this comparison, I limited by headnote number and jurisdiction. This not only limited the results to a more manageable number, but also provided a jurisdictional limit that would be critical for the needs of most researchers.[56] For this citator comparison, I focused on the numbers of citing references in common between the two citators, the numbers of unique cases for each citator, and the number of relevant cases in each set of unique results.

---

54. E-mail from Debra Myers, *supra* note 18.

55. The citation checking for this article was performed between Mar. 8, 2009 and May 8, 2009 for the first five pairs of headnotes; the rest of the research was done between July 5, 2009 and Aug. 24, 2009.

56. Generally speaking, law school is a jurisdiction-free zone. The use of multi-jurisdictional, edited cases in law students' case books masks the importance of jurisdiction, making a discussion of the importance of jurisdiction (and level of court) to a determination of the relevance of a case for articulating a legal argument a key lesson for new lawyers.

Parsing *New York Times v. Sullivan*

¶43 Below is the detailed analysis of a case comparison, again using *New York Times v. Sullivan*:

New York Times v. Sullivan, 376 U.S. 254 (1965) Westlaw headnote 4:

If allegedly libelous statements criticizing official conduct of public officers would otherwise be constitutionally protected from state court judgment awarding damages, protection is not forfeited because they were published in the form of a paid advertisement, where the advertisement was not a "commercial" advertisement but was an "editorial" advertisement which communicated information, expressed opinion, recited grievances, protested claimed abuses, and sought financial support on behalf of a movement whose existence and objectives were matters of the highest public interest and concern.

New York Times v. Sullivan, 376 U.S. 254 (1965) LexisNexis headnote 4:

If allegedly libelous statements would otherwise be constitutionally protected from judgment, they do not forfeit that protection because they were published in the form of a paid advertisement.

¶44 Shepardizing *New York Times v. Sullivan* and limiting my results by jurisdiction (California) and headnote (headnote 4), I found nine cases. Doing the same using KeyCite returned thirteen cases. Of these, four cases were in both sets of results. For purposes of this analysis, if a case showed up in both the Shepard's and the KeyCite results, it was assumed to be relevant and was excluded from the case-by-case analysis. The four cases found using both citators were:

1. *People v. Fogelson*, 577 P.2d 677 (Cal. 1978)
2. *Wirta v. Alameda-Contra Costa Transit District.*, 434 P.2d 982 (Cal. 1967)
3. *Mattel, Inc. v. Luce*, 2001 WL 1589175 (Cal. Ct. App. Dec 13, 2001)
4. *Spiritual Psychic Science Church of Truth, Inc. v. City of Azusa*, 201 Cal.Rptr. 852 (Cal. Ct. App. 1984)

¶45 There were nine cases that were only found using KeyCite. On the list below, starred cases are *not* relevant. Six of nine (67%) of these cases were relevant. Relevant was defined as a case where constitutional protection of speech that had a commercial or advertising aspect was discussed, regardless of the depth or shallowness of the discussion.

1. *Belli v. State Bar*, 519 P. 2d 575 (Cal. 1974)
2. **Imuta v. Nakano*, 59 Cal. Rptr. 2d 78 (Cal. Ct. App. 1996)
3. *Kasky v. Nike, Inc.*, 45 P.3d 243 (Cal. 2002)
4. **Bradbury v. Superior Court*, 57 Cal. Rptr. 2d 207 (Cal. Ct. App. 1996)
5. *Weaver v. Jordan*, 411 P.2d 289 (Cal. 1996)
6. **Rider v. Superior Court*, 244 Cal. Rptr. 770 (Cal. Ct. App. 1988)
7. *City of Indio v. Arroyo*, 191 Cal. Rptr. 565 (Cal. Ct. App. 1983)
8. *Welton v. City of Los Angeles*, 124 Cal. Rptr. 480 (Cal. Ct. App. 1975)
9. *Pines v. Tomson*, 206 Cal. Rptr. 866 (Cal. Ct. App. 1984)

¶46 There were five cases that were only found using Shepard's. On the list below, starred cases are *not* relevant. Only one (20%) of these cases was relevant using the same definition as was used for KeyCite.

1. **Reader's Digest Assn. v. Superior Court*, 208 Cal. Rptr. 137 (Cal. 1984)
2. **Huntley v. Public Utilities Comm'n.*, 69 Cal. Rptr. 605 (Cal. 1968)
3. **People v. Noroff*, 67 Cal. 2d 791, 63 Cal. Rptr. 575 (1967)
4. *Live Oak Publishing Co. v. Cohagan*, 286 Cal. Rptr. 198 (Cal. Ct. App. 1991)
5. **Mullins v. Brando*, 91 Cal. Rptr. 796 (Cal. Ct. App. 1970)

¶47 Thus, if a researcher were looking for additional cases involving commercial or advertising aspects of constitutionally protected speech, then, in addition to the four cases that appeared in both citators, there were fourteen unique cases to be found: nine on Westlaw and five on LexisNexis. Of these unique cases, there were seven relevant cases: six on Westlaw and one on LexisNexis. If you only used Shepard's, you would have missed 85.7% of the unique relevant cases (6/7). If you only used KeyCite, you would have missed 14% of the unique relevant cases (1/7).[57]

¶48 I performed the same comparison on the other nine pairs of headnotes. The basic information for each case and the results for the comparison of all of the pairs of headnotes are set forth in table 5. Figures 15 and 16 show the average percentage of relevant and not relevant cases found using Shepard's and KeyCite.
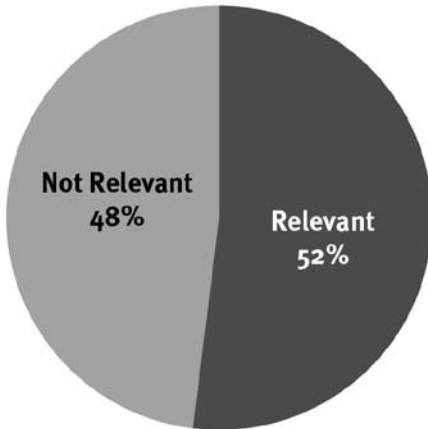
Figure 15. Relevance of Unique Cases Found with Shepard's
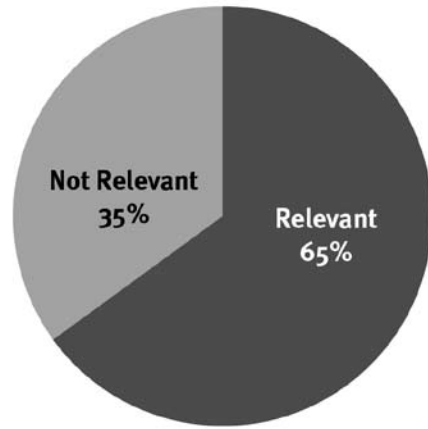
Figure 16. Relevance of Unique Cases Found with KeyCite

Using Citation Systems for Research

¶49 Once again, the percentages of unique, relevant cases might change if the pool of pairs of headnotes compared was increased. But regardless of the percentages, there are some interesting things to note about the results. Here, either the different algorithms used to match headnotes to citing references or the different

---

57. In addition to missing relevant cases, reviewing irrelevant cases wastes time. In this example, if you used Shepard's, 44% of the cases you looked at would be irrelevant (4/9). Using KeyCite, 23% of the cases you looked at would be irrelevant (3/13).

Table 5

Relevant Cases Found in Citator Comparison

Case Citation	LexisNexis Headnote	Westlaw Headnote	Jurisdiction	# only in Shepard's	# only in KeyCite	# in Both	% overlap	% Shepard's unique cases relevant	% KeyCite unique cases relevant
New York Times v. Sullivan, 376 U.S. 254 (1965)	4	4	California	5	9	4	22% (4/18)	20% (1/5)	67 % (6/9)
New York Times v. Sullivan, 376 U.S. 254 (1965)	9	9	10th Cir.	6	4	15	60% (15/25)	50% (3/6)	25% (1/4)
Brown v. Bd. of Education, 349 U.S. 294 (1955)	7	8	U.S. Supreme Court	8	5	15	53.5% (15/28)	37% (3/8)	80% (4/5)
Buckley v. Vallejo, 424 U.S. 1 (1976)	6	16, 17	U.S. Supreme Court, 10th Cir.	13	4	2	10.5% (2/19)	69% (9/13)	20% (1/4)
Escola v. Coca-Cola Bottling Co., 150 P.2d 436 (Cal. 1944)	3	2	California	1	13	14	50% (14/28)	100% (1/1)	92% (12/13)
Terry v. Ohio, 392 U.S. 1 (1968)	15	23	9th Cir.	39	12	1	2% (1/52)	46% (18/39)	77% (10/13)
Tarasoff v. Regents of the Univ. of Calif., 551 P.2d 334 (Cal. 1976)	15	34	California	2	2	1	20% (1/5)	100% (1/1)	100% (1/1)
San Diego Fed'l v. Cumis Ins. Soc'y, 208 Cal. Rptr. 494 (Cal. Ct. App. 1984)	7	4	California	14	21	3	8% (3/38)	50% (6/12)	61% (11/18)
Foley v. Interactive Data Corp., 765 P.2d 373 (Cal. 1988)	13	4	California	30	49	5	6% (5/84)	44% (11/25)	75% (33/44)
Golden Gateway Center v. Golden Gateway Tenants' Ass'n, 29 P.3d 797 (Cal. 2001)	26	3	California	7	13	7	25% (7/27)	0% (0/0)	50% (2/4)

ways in which headnotes are generated, or both, provide unique cases on both systems. It was surprising how few cases each citation system had in common; there was not that much overlap in the cases found using KeyCite with the cases found using Shepard's. As a computer programming novice I found this very surprising—it seems a simpler matter than it must be to match a pin cite from the cited case to every instance where that citation occurs. It is hard to imagine a case that cites a seed case on a specific legal topic not citing back to the original page reference. But each citation system has a large percentage of cases linking to relevant cases not found by the other citation system.[58]

¶50 Using both systems to find relevant cases will result in more thorough research. The result sets not only highlight the dangers of relying on only one system, they illustrate the benefits of redundancy in searching. Re-searching, revising, and redundancy in searching need to be highlighted for searchers who rely mainly on Google, for whom the first results may be the only results consulted.

Conclusion

¶51 Where the search process has more human intervention, it appears to deliver better results. Intermediated searching is what appears to distinguish between Westlaw and LexisNexis in the comparison of each one's classification system. Key numbers, with their intermediated correlation between the headnote and the classification topic, deliver more relevant results than any of the options offered by LexisNexis. The LexisNexis results for the Lexis Topic and "More Like This Headnote" do not completely overlap when limited by the same Focus key word limitations, although logic would indicate that there should be a fairly substantial match between the headnote and the classification topic. Both Lexis Topics and "More Like This Headnote" return result sets that include cases that do not make any reference to the targeted headnote language, suggesting that the algorithms in use have not been completely effective in making that necessary direct link between the headnote language and the classification topic.

¶52 Although the LexisNexis results are, for the sample in this article, less relevant than the West results, LexisNexis result sets do include relevant results not captured in West's Digest System. Each researcher will have to determine the best use of each system for individual research projects. Comparing the results for West key numbers, "More Like This Headnote," and Lexis Topics is one way for novice researchers to visualize the value-added benefit human indexing brings to modern searching.

¶53 The human advantage of intermediated searching does not appear to extend to citation systems. Where each citation system relies on its own algorithms, as appears to be the case for Shepard's and KeyCite, each generates a fairly unique set of results. The lack of a significant overlap for the cases in the result sets for KeyCite and Shepard's illustrates an essential problem of algorithmic searching: no one algorithm will give you all of the relevant results. That too is an important lesson for researchers.

---

58. While these unique cases did not appear in the headnote number- and jurisdiction-limited citator results of the other system, checking the unlimited citatory results was beyond the scope of this article.