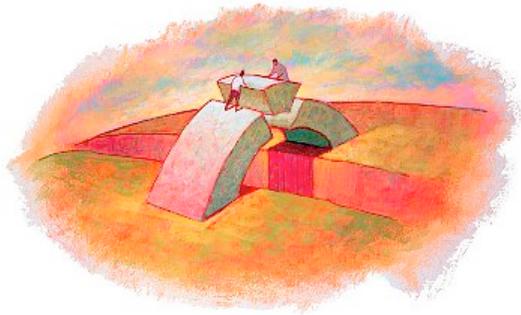


CHAPTER 5: RESEARCH: TECHNIQUES & RESOURCES

THE KEYSTONE APPROACH

The curved sides of an arch meet at the top where a wedge-shaped stone sits. This final piece -- called the "keystone" -- locks all the other stones into position. It is the central support for the



entire structure. Without it, everything would come tumbling down. When constructing an argument, researchers need to build a strong central support for their arguments by identifying the keystones for the most powerful pieces of evidence that reinforce the claim of the paper.

Logos: identifying a keystone document

Often, a single article, book, or website can be the keystone that helps organize the other important information about a topic. Sometimes these are obvious: Darwin's *Origin of the Species* is the central historical source for evolutionary studies. The book established the key terms and concepts that produced modern biology. Darwin's book was published in 1859, so the history of references to its discoveries is huge. Good students use these central texts to find more information for the topic that they are researching.

Using a keystone text is fairly simple. First, such works contain a works cited that reviews key documents for understanding the topic. These items are the network that connects to the keystone. The expert who wrote the book names this network in order to strengthen its logos (logic, evidence). One way to make certain that the text really is a keystone document is to see how often others have used it. For example, if we go to a specialized search engine such as scholar.google.com and search for *The Filter Bubble: What the Internet is hiding from you*, this is what appears:

[CITATION] The filter bubble: What the Internet is hiding from you
E Pariser - 2011 - Penguin UK
☆ 57 Cited by 3854 Related articles All 8 versions ⇨

The book has been used in 3,854 related publications. This is strong evidence that the book is important and might be considered a keystone text for the topic. These numbers imply that its content has drawn responses from others.

But suppose that you are interested in stem cells, and a textbook mentions Hwang Woo Suk's "Evidence of a pluripotent human embryonic stem cell line derived from a cloned blastocyst." You turn to scholar.google.com and find the following:

Evidence of a pluripotent human embryonic stem cell line derived from a cloned blastocyst

WS Hwang, YJ Ryu, JH Park, ES Park, EG Lee, JM Koo... - Science, 2004 - sciencemag.org

Abstract Somatic cell nuclear transfer (SCNT) technology has recently been used to generate animals with a common genetic composition. In this study, we report the derivation of a pluripotent embryonic stem (ES) cell line (SCNT-hES-1) from a cloned human ...

Cited by 724 [Related articles](#) [All 55 versions](#) [Cite](#)

It seems as if you've discovered a keystone document: 724 citations seems impressive. On the other hand, your biology professor then tells you that the article is a famous example of scientific fraud. This requires that you not only take notes, but also that you read carefully, and that you speak with the experts in the field (your professor) who can explain that many of those citations are about the fact that the work is fraudulent. Research is a social, human activity, and seeing your professors as experts reflects that fact.

Professors help direct you to keystone sources, but there are other tools that help identify such works. Such work is usually published in a peer reviewed journal or a university press.

MANAGING THE INFORMATION

Almost all your research will probably be done online. You will skim abstracts, introductions, articles, blog posts, and dozens of other types of sources. The result is often an avalanche of information that needs to be organized. Students frequently lose track of important

information, a risk sometimes controlled by relying on only a few sources. As an organizational trick, it makes sense; however, as a means of understanding the discourse, it's harmful. These five tools — browser add-ons and web sites - help you control the flow of information produced by your searches. Many other tools perform similar functions:

1. www.mendeley.com Mendely's most useful function is its ability to manage references. This function enables users to share research. It's a bibliography manager, an efficient and easy way to create and share bibliographies. endnote.com and zotero.com are similarly excellent.
2. www.unpaywall.org enables you to legally avoid the paywalls that keep students from reading millions of important articles. Other add-ons that perform the same function are <https://openaccessbutton.org/>, and <https://core.ac.uk/services/discovery/>
3. scholarcy.com summarizes long .pdf articles and highlights key terms. Note that scholarcy.com only works with Chrome.

TECHNIQUES FOR FINDING EXPERTISE

Good searches require good questions. Good questions contain precise terms. Thus, the task seems simple: use good terms to search for reliable information. Of course, it is not that simple because you cannot know what terms are useful until you understand your topic. It is no surprise that students feel paralyzed early in the search process. Writers must recognize that the identification of key terms is a two-steps-forward, one-step-back process.

What are good places to build key terms for searches?

The list of key terms is incomplete at the beginning of your search. As writers read and take notes, they find new terms that can be used in additional searches. Good students keep a separate, growing list of key terms drawn from the following sources:

1. *Class notes*: technical terms the professor has used or emphasized in class
2. *Textbooks*: words emphasized with italics or bold print because they are shared by experts in the field. Read the table of contents to find key terms. Read the index to find key terms.
3. *Titles of articles*: professors often provide important articles with bibliographies that identify key items in the discourse.
4. *Conversations with professors*: professors have lifelong interests in the topic they teach. Asking about key terms is evidence of being a serious student.

Identifying Expert Terms and Sources With Online Tools

1. Use specialized search engines, scholar.google., etc. to find overviews — called “review articles” — of the topic.
2. Recognize documents that provide an expert overview of a topic; note that full text documents are not always necessary at this stage because writers need terms and concepts.
3. Identify key scholars and follow them on social media such as Twitter and their blogs to access important questions and current developments in the field.

Some of the most interesting topics are new. For example, Eli Pariser’s book about “the filter bubble” is an example of a topic that has become influential after its publication in 2011. Thus, scholarship is not yet fully developed. However, many important scholars are writing about it in blogs, on Twitter, and in other digital publications. What they say is crucial to creating a good paper about the topic. Many scholars use Twitter to identify significant postings, new articles, and other material. The author of *The Filter Bubble: How the New Personalized Web Is Changing*

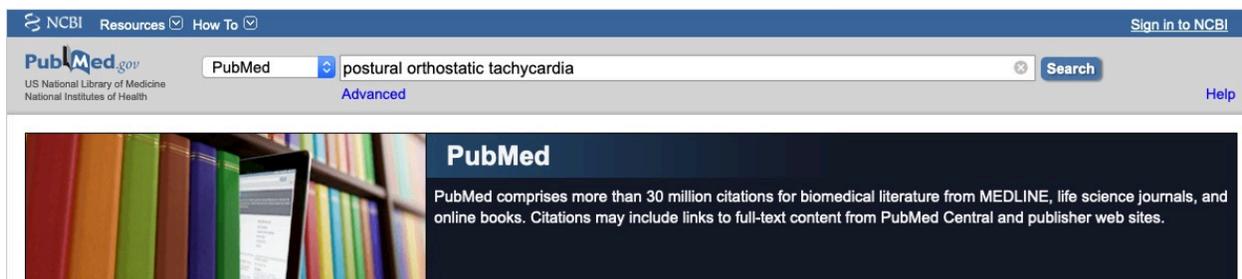
What We Read and How We Think is Eli Pariser. Through his Twitter account, he “follows” 733 authorities on the topic. Anyone interested in Pariser’s work can log in to Twitter and see the discourse community of which he is a part.

A professor whose class was writing about the filter bubble created a Twitter account for her students, and then emailed this message to the class:

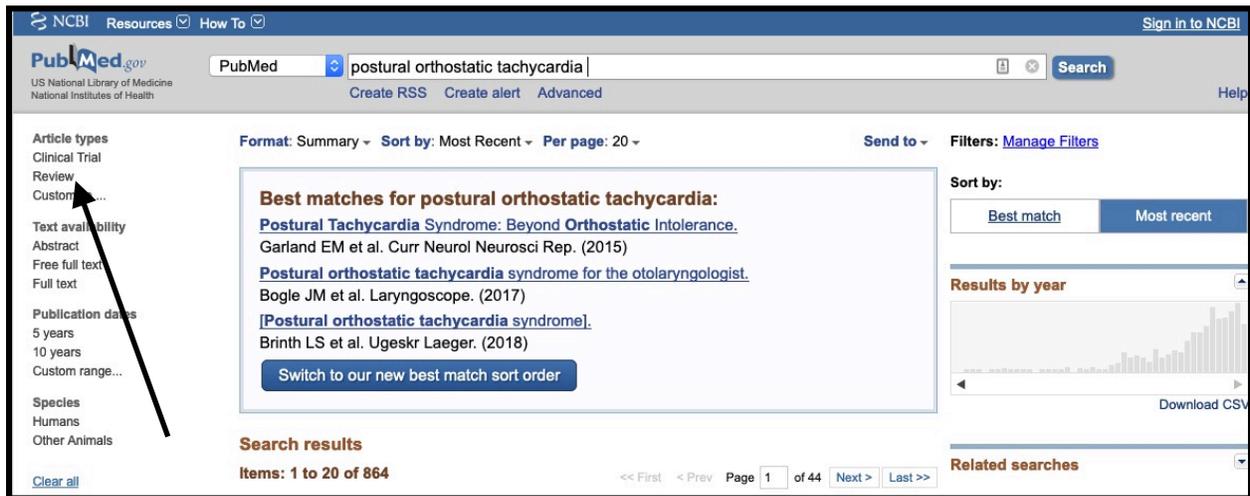
For Twitter, think of it as a way to stay up to date on the kinds of things that were discussed in the Pariser talk and the essays we read. Too often, students are taught that research means looking up a few things “on” google and then they are done. This is not the way we should be doing things. We need to be aware of the on-going conversations about these ideas--in our case: privacy, filtering, algorithms....and be aware of how people are talking about them and thinking about them. We have talked about developing a “working knowledge” of a topic, and Twitter is one way to do that. If you take a look at what has been tweeted since class began, there is lots of stuff that pertains to what you are writing about and could serve as excellent resources for your paper. Lots of experts in various fields are on Twitter and they use it to communicate with people who are interested in the same things. So, if you find any relevant essays, you should tweet the link to me and I will send it out to the class.

Again, knowledge is social, and social media can strengthen that knowledge.

Databases often categorize articles, and offer a link to review articles. For example, nurses, doctors, and other medical specialists rely on PubMed (<https://www.ncbi.nlm.nih.gov/pubmed/>). It is the National Library of Medicine at the Federal Government’s National Institute of Health. Let’s assume that you go to the site to search for informational about “Postural Orthostatic Tachycardia.” You would be offered the following:



By clicking on “search,” you’d be offered this:



The page is much longer than what’s shown here. It includes articles, a list of related search terms, and other valuable ways to focus your search. For someone new to the topic, the most valuable link is the “Review” link which offers the articles that provide an introduction to the topic. Look for the link to review articles, or use the term in your searches.

Many journal articles begin with an “abstract,” a brief summary of the article’s method, evidence, and claims. By providing the key terms of the document, the abstract helps guide your search for information. In the early stages of a search, students can benefit from reading only the abstract. Book introductions are like the abstracts for articles because they present the book’s main claims and major evidence. As with abstracts, they orient students to a field and supply important terms for further searches. Both abstracts and introductions summarize a single work. Many book reviews examine groups of books that deal with the same issue. Many book reviews comment on general problems in a field and thus help introduce students to a new field. Some articles look back over previous research in a field and review its strengths, weaknesses, and value for future discovery. Such articles are called “review articles,” and they are valuable for surveying the field. Often, their works cited provides a treasure trove of reference sources.

TOOLS FOR FINDING INFORMATION

Search engines can find the sources needed to create a map of the existing information about a topic. While most students have “googled” information, few know how to select the right search engine, identify key terms, or do an advanced search. Knowing how to do these things saves time, and it also produces the useful results needed for a good paper.

What is a search engine?

Search engines are the librarians of the Internet. When you want to know about a subject, you ask the search engine where it is stored; then, it tells you how to get there. The search engine does *not* have the information. It tells where on the Internet that the material exists. Thus, no one ever “found it on Google.” startpage.com, duckduckgo.com, google.com, bing.com, yahoo.com, and other general search engines provide a list of links to other web sites. Like a librarian, the search engine might not know the subject, but it does know where to find information about it. Remember: a search engine *does* something; it finds things. Where does it find them? It finds them in an index of web sites, words and images. Sometimes the index includes materials that have been reviewed by experts. Sometimes the index comes from a wide range of sources, some of which are expert and others of which are not. It is important to know what the search engine has indexed and can provide to you.

So, let's correct the analogy between search engines and librarians: real librarians know how to lead you to key terms, key disciplinary discourses, key databases, and all the other subtleties that underlie a strong paper. They are experts who understand the nature of credibility, discourse, and discourse communities. Search engines perform their function after the intellectual work of librarians maps out an information strategy.

What kinds of information do general search engines locate?

About 85% of all English language searches rely on Google, Bing, or Yahoo. These general search engines provide a map to other places with information related to the key words in the search. Note that different people will get different results even if they both use the same terms. Many websites and search engines record a user's interests, past online choices, geographic location, and other factors. These factors are used to "personalize" the results of the search. In some ways, this is good. If a user searches for pizza, it helps that only pizza parlors in her/his area show up on the list. On the other hand, the results for politics, ideology, beliefs, and values are likely to be tilted toward those that agree with what the searcher already thinks.

Personalization can turn a search engine into a kind of sneaky librarian who only sends you to information s/he wants you to know. Be suspicious of search results. The good news is that some general search engines, for example, DuckDuckGo.com and startpage.com do not personalize results.

Do search engines search the entire web?

Even the most powerful search engine indexes only the surface of the web. The "deep web" escapes indexing, and it is probably 100 to 1,000 times the size of the surface web. Other search engines appear *after* arriving at a web site. They search only within the site. For example, many students turn to Purdue University's splendid Online Writing Lab (OWL) at owl.english.purdue.edu. This site has a search engine, but it *only* searches inside the Purdue database of materials. This assures users that the information has been created and approved by the Purdue writing program. Many businesses use a similar type of restriction so that customers stay within their products.

How are the results of general searches organized?

The results of general search engines such as Bing, Google, and Yahoo appear on the page according to formulas (algorithms) that reflect *your* history of searches, *your* history of links

followed, the number of links to a site, and many many many other factors. The results are not ranked in order of importance or quality. Be leery of using *general* search engines for academic purposes. Note that the results from specialized search engines are controlled by different factors (see below).

Are there specialized search engines?

Other search engines search specific types of publications. They give quick access to expert information. The following engines limit themselves to scholarly publications and are available without charge. Both engines listed below give links to related articles and to other works that cite an article.

- www.scholar.google.com
 - For scholarly work in a variety of disciplines. Includes conference proceedings, books, and journals. Provides citation information. Helps focus searches, identify key journals, and key terms.
 - Presents links in an order influenced by the usual behavior of scholars: author, publisher, citation frequency, and language.
 - Frequently cannot provide full texts of articles and other materials, but does make using the library databases easier.
- <https://citeseerx.ist.psu.edu/index>
 - For scientific information on both the surface web and much of the “invisible web” including pdfs, conferences papers, etc.

By identifying scholarly sources, these sites become key tools in students’ research.

The Filter Bubble

Many search engines record what each user does online. Even if a searcher has not logged in, the search engine still knows many things that determine the results provided. This information includes the kind of computer system used, the browser, the location of the user, and probably 50-100 other facts. A searcher who signs in reveals much more. In either case, the result is both

wonderful and dangerous. It is wonderful because the results are “personalized.” For example, if you are a poor student looking for a hotel in New York, it will not help you to be offered the rate for a room at The Plaza. Information about a Motel 8 in New Jersey might be much more useful. The search engine filters the results they send by keeping track of your other searches. Do you search inexpensive department stores for clothes, or do you search Bergdorf Goodman? Did you check prices on used cars or on Cadillacs? Your history tells about who you are and what you are likely to buy. This information makes you valuable because it lets the search engine make a promise to advertisers. It promises them that it will advertise their products to the most appropriate market. Personalization reveals something important: *you* are the product being sold. Your attention to web pages is valuable to advertisers, and the search engine is keeping track of what you are likely to want so that the ads can be posted in front of the people most likely to buy.

Personalization also has risks. By keeping track of our beliefs, desires, and behavior, a search engine can become a sort of echo chamber where searchers only find material that agrees with what they already think. Liberals hear only liberal ideas; conservatives hear only conservative ideas. No one encounters ideas that challenge her or his beliefs. This is called “the filter bubble.” In the filter bubble, future information is based *only* on past choices. The future cannot escape the past. Many scholars worry the filter bubble will take away the benefits of surprise, contradiction, and argument. They fear it will harm both individuals and our culture’s need for genuine dialogue.

Discussion: The activities of Cambridge Analytica epitomize the dangers of personalization. Using the search engines mentioned above, search for the most recent information on its activities. What are some of the proposals that seek to control such risks?

Discussion: Algorithms frequently produce harmful results. Read <https://www.theguardian.com/technology/2019/oct/14/automating-poverty-algorithms-punish-poor> for an overview of how this process works. Do any of the risks pointed out in this article apply to you? How can you protect yourself? How would you search for information on protecting yourself?

The Relevance Paradox

The filter bubble seems dangerous. However, critical readers understand that it is a mixed blessing. They recognize that writers have judged some evidence important. Those selections make some ideas more important than other ideas.

Good readers also try to recognize what has been ignored or not taken seriously. While good readers know that some things will be consciously “unknown” in the reading, they also know that there are “unknown unknowns” in any explanation. Secretary of Defense Donald Rumsfeld infamously said:

Reports that say that something hasn't happened are always interesting to me, because as we know, there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns -- the ones we don't know we don't know.

Rumsfeld’s statement maps out the risks that face any critical reader. These highly skilled readers use their recognition of unknown unknowns to insist that even the best critical reading is unfinished.

LIBRARY DATABASES AND WHY WE SHOULD USE THEM

Library databases are collections of filtered information. They often contain information that is not available through other search engines. This information is organized to allow easy access to complete articles and other materials. Some organizations and institutions have built databases filled with carefully filtered information. For example, JSTOR contains all the articles from more than 1,000 journals. It also contains related materials. JSTOR relies on established experts — on the discourse community — to judge whether or not something should be included. Factors

such as historical significance, citation patterns, subscribers, and recommendations from libraries and scholars determine if a journal or other source will be included. Almost all have a search engine within them that finds material in the database.

Question: Search engines such as scholar.google.com locate important articles. How is a search engine different from a database such as JSTOR?

Do some databases focus on specific disciplines?

Many databases cover several fields, but some are highly specific:

- Art: JSTOR, Project Muse, Art Full Text
- Business: Business Source Premier, Central Business File ASAP, Standard and Poor's Net Advantage
- Computer Science: Computer Source, INSPEC
- Education: ERIC, Educator's Reference Complete
- Engineering: Compendex, INSPEC
- History: Historical Abstracts, JSTOR
- Literature and Language: MLA International Bibliography, JSTOR, Voice of the Shuttle
- Mathematics: MathSciNet (<https://mathscinet.ams.org/mathscinet/>)
- Medical Information: Pubmed (<https://www.ncbi.nlm.nih.gov/pubmed/>)

Again, the guidance of librarians enriches your thinking by pointing you to databases, bibliographies, government documents, and myriad other sources that focus on your topic.

Are library databases the best sources of information?

Major databases act as specialist librarians who know what is recognized as most important about a subject. Using library databases puts you inside the boundaries of a discourse community. Not only do they filter out questionable content, but also they connect to the

network of experts, institutions, publications, and ideas that define the field you are researching.

The information in library databases represents the careful judgment of experts. Such careful judgment takes time. Thus, there's often a lag between research and publication. This lag reflects the nature of building knowledge. Nonetheless, some important issues erupt quickly and receive attention outside peer reviewed publications. Major scholars are likely to write about them in publications that are not part of library databases. These insights, discussions, and critiques appear in blogs, topical web sites, and "middle brow publications." Middle brow publications present detailed information and background to audiences who are well-educated but not specialists. They do not use peer review, but the authors are highly qualified. Thus, you must recognize that good research reflects a wide range of information gathering. Good research uses Twitter, library databases, Facebook, scirus.com, Bing, DuckDuckGo.com and other tools that keep you in touch with the dynamic nature of knowledge.

OTHER DATABASES AND ONLINE RESOURCES

Government databases cover many subjects. The sites end with .gov to signal the source of information and control. A few of the many valuable sites include the following:

- Bureau of Labor Statistics: www.bls.gov
- Census Bureau: www.census.gov
- Centers for Disease Control and Prevention: www.cdc.gov
- Childstats.gov: www.childstats.gov
- Fedstats: www.fedstats.gov
- National Archives: www.archives.gov
- National Center for Education Statistics: www.nces.ed.gov
- National Institutes for Health: www.nih.gov

Many other government web sites provide valuable information.

Collections of complete books, articles, and other materials are available at the following sites:

- The Internet Archive: www.archive.org
- Online Books: online books.library.upenn.edu
- OpenLibrary: openlibrary.org
- Project Gutenberg: www.gutenberg.org
- Scholars' Lab Digital Resources: www.lib.virginia.edu/scholarslab/resources
- ScribD: www.scribd.com
- Google Books: books.google.com

Advanced search techniques

No one lives long enough to read the millions of results that come from a poorly designed search. An effective search depends not only on using precise terminology, but also on knowing some simple rules for getting useful results. Three simple terms produce more useful results: *AND*, *OR*, and *NOT*. Consider the following example: suppose you want to search for information about freshwater clams in Lake Erie:

*Using quotation marks, AND, +, and blank spaces to **narrow the result** of searches*

- Entering **freshwater clams** produces about 1.6 million hits on Google.com. The search finds websites with both words, but not necessarily in that order. The blank space between the words means “both words but not necessarily in that order or even next to each other.”
- Enter **“freshwater clams”** and the results drop to about 45,000 hits. The quotation marks narrow the search to a specific phrase. The search finds websites with only the phrase “freshwater clams” in exactly that order.
- Entering **“freshwater clams” AND “Lake Erie”** produces only 4,500 hits because the *AND* combines with two phrases in quotation marks. The search narrows to *only* sites with both phrases exactly as written in the search bar. Some search engines require a + sign or the word *AND* between words to narrow the results of a search. Google assumes that a blank space means “*AND*.”
- Using *OR* to **broaden the result of** searches by including any appearance of any term.

When different terms refer to a similar concept (school, college, university), a good search may need to find sources for all three. The term OR enables a search engine to combine all three and provide more results. For example, a search to determine the average number of computers in all schools might need to use OR. Searching separately on Google gives the following results:

- Entering school produces 700 million hits
- Entering college produces 500 million hits
- Entering university produces 400 million hits

The total hits for the individual searches is 1,600 million hits. These searches can be combined with “OR.”

- A single search using school OR college OR university 1,000 million hits.

Note that the sum of the individual searches is larger than the single search using OR. “OR” works on two levels: first, it increases the number of hits by adding together the results for all the terms; second, it eliminates the duplicates of the separate searches.

Using NOT to narrow the result of searches

Some searches ask for information about a category, for example, fish. The search seeks information about the category but does NOT want information about trout. Using NOT can solve this problem:

- Entering fish produces 656 million hits on Bing.com
- Entering trout produces 84 million hits on Bing.com
- Entering fish NOT trout produces 376 million hits on Bing.com.

Combining advanced search terms

AND, NOT, and OR can be combined in many ways. Experimenting with their combinations can produce highly specialized results. Information about such searches is available through a variety of sources including librarians, handbooks, and web sites.

BUILDING ON THE TEQ SHEET: ANNOTATED BIBLIOGRAPHIES

TEQ Sheets have three parts that lay a foundation for our note-taking. It creates an overview of the document that can be expanded by adding detail. These enlarged notes are called “annotated bibliographies.”

Annotated Bibliographies collect additional information about a reading by asking three basic questions:

- *Summary*: What is the most important information presented in the reading? This is a factual summary that does not include your responses to the information.
- *Evaluation*: What level of credibility does the source bring to discussions of the subject? Is it peer reviewed? What expertise can the author claim? Do the citations place the article on a map of clearly understood and accepted discoveries?
- *Integration*: How does the source sharpen your own map of the topic? Does it sharpen your idea by providing additional information? Does it sharpen your detailed additions to the map by challenging what you would like to say?

Paraphrase and Quotation

Good writers identify key sentences, facts, and phrases from important sources. Showing the source of such material strengthens a paper. Writers have two major tools for including the supporting network in a document:

Quotation

Quotations use the *exact* phrasing of a source. The language is identical, and the material is placed inside quotation marks. The quotation’s source must be named.

Paraphrase

Paraphrases put the materials of others in new words. The revised language keeps key concepts. Often, the phrasing integrates the source with the style and voice of the paper that uses it. Like quotation, paraphrase requires that the source be named. Failing to provide a citation is a form of plagiarism.

Note: some students have been told or believe that by changing a word or two in each sentence, there is no need for a citation. This mistaken notion can lead to charges of plagiarism. Don't do it.

WHY ACADEMIC PAPERS CITE SOURCES

Readers expect to see the connections to the existing map of the topic. They expect that writers will identify those connections in the text of the paper. The references in the text of the paper will enable readers to find complete information about the sources in a list (the “works cited”) at the end of the paper. Writers show that their work connects to expert knowledge by citing sources. A writer who names these links in the document gains credibility. A writer gains credibility by naming the network of experts who influence the document. Citations have other advantages:

- they recognize the history of the topic
- they reveal the various ideas and positions that frame a topic
- they highlight the difference between the writer’s own insights and the insights to which it connects

Naming sources through careful citations is so important that expert writers use standardized formats that make it easy to retrace the sources for a work and understand how the new document adds to the understanding of a topic. Most U. S. professors expect that you will follow one of two systems:

- Modern Language Association (MLA): for the humanities

- American Psychological Association (APA): for psychology, sociology, business, nursing, economics, and other field

These are the two dominant styles, but others such as Chicago style are also important. Use the format style required by the publisher, professor, or employer. These formats are part of the formal structure of an academic paper. See Chapter 7: "Structuring the Document."