The Essential Guide to Using the Web for Research

Finding High-Quality Information

Contributors: Nigel Ford
Book Title: The Essential Guide to Using the Web for Research
Chapter Title: "Finding High-Quality Information"
Pub. Date: 2012
Access Date: February 05, 2014
Publishing Company: SAGE Publications Ltd
City: London
Print ISBN: 9780857023650
Online ISBN: 9781446287927
DOI: http://dx.doi.org/10.4135/9781446287927.n4
Print pages: 37-53

http://dx.doi.org/10.4135/9781446287927.n4
[p. 37 ↓]

Chapter Four: Finding High-Quality Information

Why You Need to Know This

- Whatever the type of assignment on which you are engaged – whether a basic level essay or an advanced research project – you need to find and make use of authoritative information.
- Some of this information may be provided for you by your lecturers, for example in the form of recommended reading lists. However, as you progress from basic to more advanced levels of study, you will increasingly be expected to be able to find high-quality (relevant and authoritative) information for yourself.
- First, you need to know exactly what is meant by 'authoritative' information. Those marking your work will expect you to know this, and to avoid the use of information that is not appropriately authoritative.
- You also need to know how to find it. Whilst it is easy to find large volumes of information, it is more difficult to find the type of high-quality information you require for your academic work.
- You need to be aware of the problems and potential pitfalls of information seeking so that you can avoid and overcome them. This chapter explores the nature of such difficulties. Tools, techniques and strategies for countering them will be introduced in Chapters 5 to 8.

This chapter will define what is meant by 'authoritative' information, and will explore the processes entailed in effectively finding it. But before we go any further…

Think

What exactly is *information* – and how does it relate to knowledge and learning?

[p. 38 ↓] As is the case with the terms *understanding, knowledge* and *learning* presented in Chapter 2, it is important to note that *information* can be defined in many ways, for different purposes, and from different perspectives.

The following is the working definition that I have adopted for this book. As stressed in Chapter 2, the important thing is not to attempt to present some universal definition of these terms upon which everyone can agree (an impossible task), but to ensure that both you as reader and I as writer have a shared understanding of how they are used in this book. Here, then, is my working definition of *information*, and how it relates to *knowledge* and *learning.*

Information is the expression of (a part of) someone's knowledge. Information may be expressed in recorded form (e.g. a book or video) or communicated live (e.g. a lecture or conversation).

If learning is the process of acquiring new knowledge, then information is the raw material fuelling this process.

Information is transformed into knowledge via the process of learning.

There is a universe of information sources ‘out there’ accessible to you via the web, many of which are freely available. Others require payment. However, your university will have subscriptions to a wide range of such sources as well as its own library collection of resources. A key skill required of you is to be able to effectively navigate the range of information sources available to you both from your library and freely over the web. We will be examining the range of different types of information source available to you later in this chapter. There is also a range of different types of search tool that can enable you to find these different types of information. Again, we will be exploring these later in this chapter.

But first, let us explore the notion of *authority* in relation to information. You will recall from the discussion of evidence-based claims and arguments in Chapter 2 that *authoritative* sources should be used to support assertions and propositions that you present in your work. In the discussion of *critical thinking* in the same chapter, we noted the need to distinguish between *authoritative* and *non-authoritative* sources.

Defining ‘Authoritative’ Information

At all levels, from basic understanding of a new topic to the critical analysis of controversial issues, you need to make use of authoritative information. This is necessary since the person marking your essay or project will be asking:

> **[p. 39 ↓]** *What evidence have you put forward in this essay/dissertation to convince me that what you are saying is reasonable?*

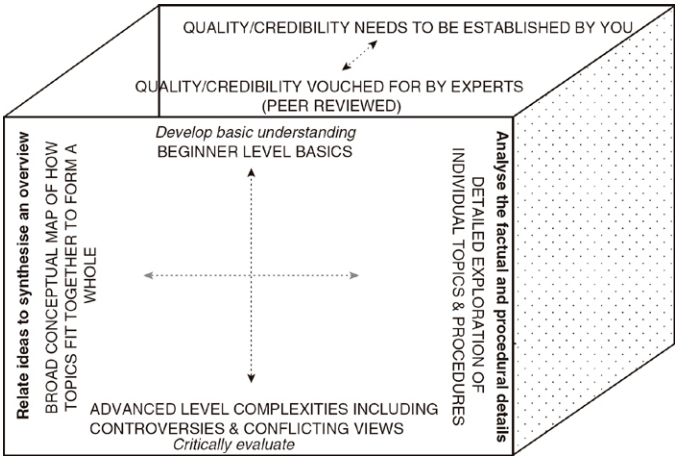
There are plenty of information sources, on every subject under the sun, written by well-meaning (and sometimes not so well-meaning) people. However, if you don't know anything about them it is perfectly possible that the information is flawed. It may be poorly argued or it may be partially or completely inaccurate.

But how do you know if information is of high quality, especially if you are new to the topic? Well, in the academic world a key concept relating to quality is *peer review.* Peer-reviewed information is information that has the stamp of credibility because it has been subjected to *independent review by experts.*

Books, journals and conference proceedings published by scholarly publishers or professional organisations will generally be peer reviewed. This means that before a book, book chapter, journal article or conference paper is accepted for publication, it will be sent to independent reviewers who are expert in the topic area. Frequently these reviewers will recommend or require certain changes to be made before the work is accepted. This is a way of ensuring that high academic standards are met. Theses and dissertations produced for an academic qualification are also peer reviewed in that they are examined by independent examiners chosen for their relevant expertise.

Figure 4.1 integrates the notions of *quality* and *credibility* relating to information sources into the model previously presented. These appear as the new dimension labelled on the top face of the figure.

Figure 4.1 Credibility of information sources



[p. 40 ↓] Finding and using peer-reviewed information is important in relation to your study at both the top and the bottom of Figure 4.1 – i.e. whether you are just beginning a new topic/subject or are at a more advanced stage.

In the early stages of your learning about a new topic or subject, you are trying to discover what are generally agreed by experts in the field to be the key facts, ideas and arguments making up the topic or subject. At both basic and advanced levels, it is important that what you say is based on authoritative sources. In other words, you need to back up what you say in your written work by referring to these sources. By choosing those that are peer reviewed you can be assured of a certain level of academic quality.

At a basic level, you should refer to these sources to support your claims and arguments. Recall the example given in Chapter 2 of the quality difference between the two claims:

> ‘Females tend to be better at expressing positive emotions when using social networking sites.’

> ‘Females tend to be better at expressing positive emotions when using social networking sites (Thelwall, Wilkinson and Uppal, 2010).’

where the paper by Thelwall, Wilkinson and Uppal was a peer-reviewed source.

At more advanced levels you will probably come across a range of different arguments, views and perspectives on your topic. At this level, you will be expected not just to present these, along with references to where you found them but also to assess them – to compare and contrast them in terms of their contribution and limitations in relation to answering your essay or project question. You will also need to present your own considered view, made in the light of this possibly conflicting evidence.

At this level, being peer reviewed does not indicate that an argument or view is necessarily ‘correct’, but rather that it is reasonable – i.e. evidence-based. You can still find diversity of opinion in peer-reviewed sources. You may disagree with a particular view, but to argue against it you will need to present counter-arguments and appropriate evidence. You may find such counter-arguments and evidence in other peer-reviewed sources.

In Chapters 4 to 7 we will explore how to use the web to find information that has the authority stamp of peer review. However, non-peer-reviewed sources, such as blogs or websites of individuals and organisations, can at times be useful – but with appropriate caution and evaluation. When you are using sources that are not peer reviewed, you will need to establish for yourself their credibility. You will also need to establish and justify to your reader the credibility of these sources and the reasons why you have used them rather than a more conventional peer-reviewed source.

So why, indeed, would you want to use a non-peer-reviewed source? Well, knowledge generally progresses via the interplay between old and new ideas. New claims, arguments, evidence and perspectives emerge that may challenge as well as build on previous ones – even those previously well established and widely regarded as authoritative. However, leading-edge ideas may not always be readily available in **[p. 41 ↓]** traditional peer-reviewed sources. This may be due to the lead-in time required to get an article reviewed and published in a peer-reviewed journal. Also, it may sometimes be relatively difficult to get a highly controversial article published in a top

journal, especially if the article runs counter to accepted wisdom in terms of content or approach.

Thus, there is a constant tension between consensus and novelty in the generation of new knowledge. Creativity entails forging links between concepts in previously unthought of ways. Extreme consensus-based thinking (strongly emphasising convergent thought processes), in which you only ever use ideas that have been thoroughly tried and tested, will at best result in the slow, steady accumulation of knowledge, but is likely not to result in the discovery of really new ideas and directions. Extreme novelty-based thinking (emphasising divergent thought processes) may spark new ideas and directions, but at the risk of generating instead ideas that are so ‘off the wall’ that they have little value.

This is reminiscent of Pask's learning styles. Recall that procedure building unconstrained by complementary description building led to fragmented learning – failing to see the wood for the trees. Description building unconstrained by complementary procedure building led to over-generalisation – essentially similar to overly divergent thinking in which concepts are linked in previously unthought of, but ultimately spurious, ways.

An important question is to what extent should you…

- ‘play safe’ in your essays and projects, and keep strictly to peer-reviewed authority sources and well-established information; or
- be creative and try to forge new connections between ideas?

In relation to research projects where you are gathering and analysing your own data, you can generate genuinely new knowledge. However, you are not normally expected to come up with genuinely new knowledge in more basic essays. Rather, you are expected to show mastery of the subject matter in terms of being aware of the key concepts and issues, and to show that not only have you understood and analysed them, but that you have thought critically about them and formed your own view based on the available evidence.

So my advice is – be cautious in using non-peer-reviewed sources. If you feel that they would contribute usefully to your essay or project, you should first evaluate their credibility (see Chapter 9 for details of evaluation criteria and how to apply them), then consult your lecturer before proceeding.

Information Seeking

There may be times when a quick and easy search using a general search engine like *Google* is perfectly adequate and retrieves just the information you need. The major search engines have developed sophisticated techniques to enable quick and easy searches to retrieve much useful information.

[p. 42 ↓]

Some Basic Problems

However, as we saw in the previous section, *Google* and the other main search engines are not the only tools you can use to find information. Nor are they always the best. Much depends on the nature of the information you need. It is all too easy to get into the *Google* habit and almost automatically use it for all yours searches. However, depending on what you are searching for, there can be problems with this.

Search Tools

Think

There is a place somewhere in the world called ‘Place’. Use the web to find out where it is. [Starting in *Google*, this information can be reached within four clicks.]

‘Place’ is located… [where?]

Searching using *only* a search engine such as *Google, Yahoo, Bing*, etc., is likely to prove extremely difficult in a case like this. You should bear in mind that, although this particular example may seem rather contrived, it is illustrative of a very common problem that can affect your searching.

Most documents indexed by the search engines using the word ‘place’ are about places – not specifically places called ‘Place’. Although documents that *do* talk about the place ‘Place’ will *also* be indexed, they will be completely hidden in the wealth of documents dealing with the more common meaning of the word. On the first page of hits, *Google* reports:

> ‘About 1,260,000,000 results’

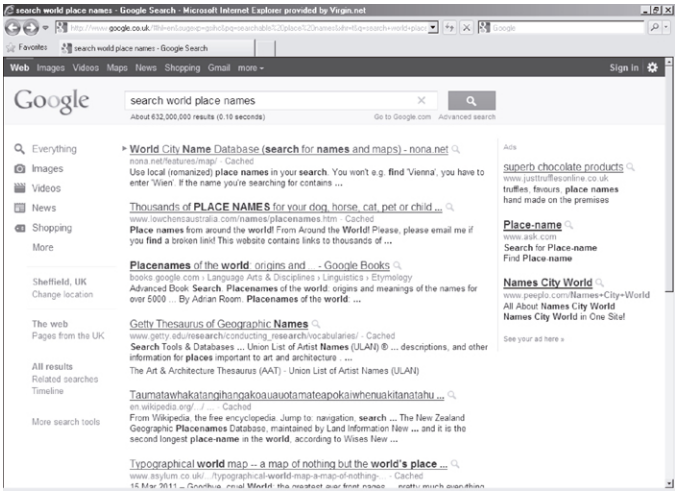
The key is knowing about and using a more appropriate search tool. In this case, you need a tool that specialises in place names. So if, rather than searching for ‘Place’ we search for, say:

> *search world place names*

we retrieve a list of sites that allow us to engage in a more specialised search by a tool that ‘knows’ that we are wanting place names. Figure 4.2 shows the results of our search.

[p. 43 ↓]

Figure 4.2 Search for search world place names in Google



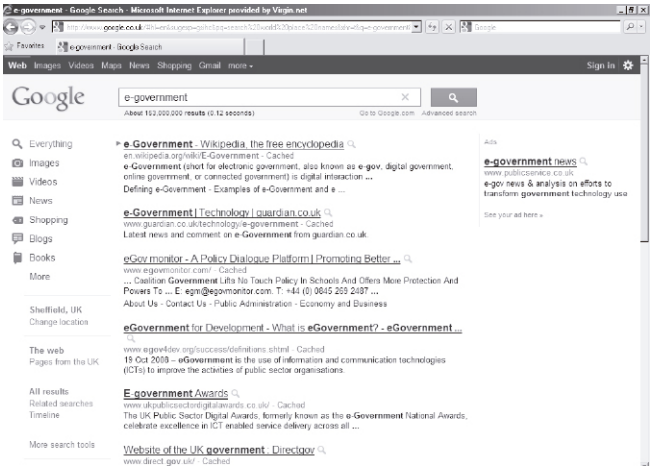
The first item on the list will take us to a site that will enable us to search specifically for place names. Search for ‘Place’ and we will discover exactly where it is.

As noted above, this rather extreme example illustrates a very common and pervasive problem – namely, that the precise information that you want may be hidden by less relevant information. This happens when you can't tell a search tool *exactly* what it is that you need. You can tell *Google* that you want information on ‘Place’, but you can't specify that you only want to use it as a *proper noun* and not in the more general sense.

Similarly, you might want to search for information on, say, *e-government* using *Google*, but you can't specify that you only want high-quality authoritative information suitable for an academic essay. You certainly get an interesting list of sources (see Figure 4.3), but you can't limit the hits only to high-quality academic sources. You end up with a mixed bag in terms of quality. Again, it is a case of not being able to sufficiently and accurately specify to *Google* exactly what you need.

As shown in Figure 4.3, *Google* has found some interesting sources relating to *e-government*, but they are by no means all suitable for use in an academic essay or dissertation.

Figure 4.3 Google search results for e-government



[p. 44 ↓]

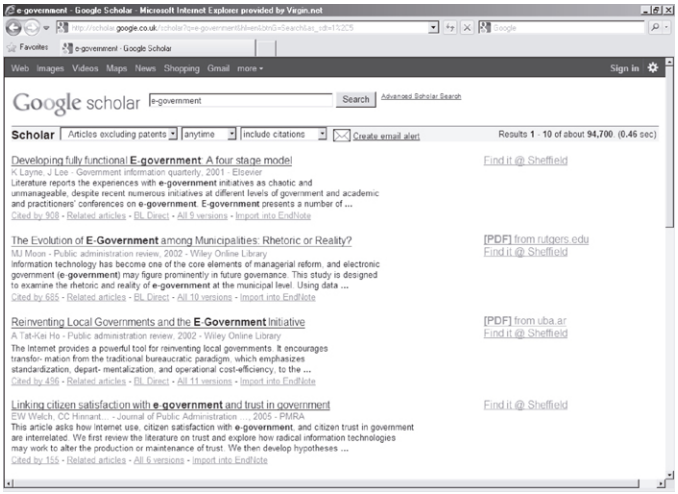
Think

Can you think of a way to search for *e-government* using *Google* – but in such a way as to get only good-quality academic sources?

If you thought that you should use some other tool – you're right. There are a number of tools designed to help you find high-quality academic sources. Some, like *Web of Knowledge* and *SciVerse Scopus*, are subscription-based, but your university will provide access to these or equivalent tools. *Google Scholar*, however, is freely available on the web. It is, as its name suggests, a scholarly alternative to the general *Google* search engine. Search for *e-government* in *Google Scholar* and you will retrieve good quality academic sources (Figure 4.4).

[p. 45 ↓]

Figure 4.4 Google Scholar results for e-government



As we will see in Chapter 8, we can refine our searches in *Google Scholar* to specify, for example:

- The most recent information sources;
- Those within a particular discipline (Social Sciences as opposed to Environmental Science, for example);
- Those written by a particular author;
- Those written in a particular year or between certain dates;
- Those available freely to anyone online via the web;
- Those available online to members of our particular university via the web;
- Those not available online but available in our particular university library;

and more…

We will be looking in detail at a range of such tools in Chapter 8. The point here is that you need to be aware that a vital component of effective information seeking is choosing the most appropriate search tool. Before rushing to a general search engine, you should first make sure that there is not a more appropriate tool that would provide you with much more effective results.

[p. 46 ↓]

Search Strategies

Even assuming that you have chosen the most appropriate search tool to use, there are problems that can all too easily hinder effective searching. These potential problems mean that it is well worth your taking a little time to think about your *search strategy.*

When two people speak, they can easily establish the fact that they are talking about the same thing. In other words, they can negotiate any small differences in the wording they use almost without having to think about it. For example, if a student is studying effects of *poor school attendance* and his lecturer knows that there is an excellent book called *The Effects of School Truancy*, she has no hesitation in recommending it. For both lecturer and student, *poor school attendance* and *school truancy* are to all intents and purposes synonymous.

However, we cannot assume that such a translation necessarily takes place in a search engine. Let us perform two searches in *Google.* In the first, we will enter the keywords:

the effects of poor school attendance

In the second, we will enter:

the effects of school truancy

Figure 4.5 shows the results of the two searches – the first at the top and the second below it. This figure was produced by a website (http://www.thumbshots.com/Products/ThumbshotsImages/Ranking.aspx) which allows you to compare different searches in the same search engine – or indeed, the same search in different search engines. The top row of balls represents the first 60 information sources retrieved by *Google* in response to the first search (the *effects of poor school attendance)* and the bottom row shows items retrieved in response to the second search *(the effects of school truancy).* Each ball represents an information source. The balls representing any *common* information sources (i.e. ones retrieved in *both* searches) would be indicated by a line linking them across the top and bottom rows (an example of this is shown in Figure 4.6).

Figure 4.5 Comparison of Google results for two searches using only a slightly different way of expressing essentially the same topic

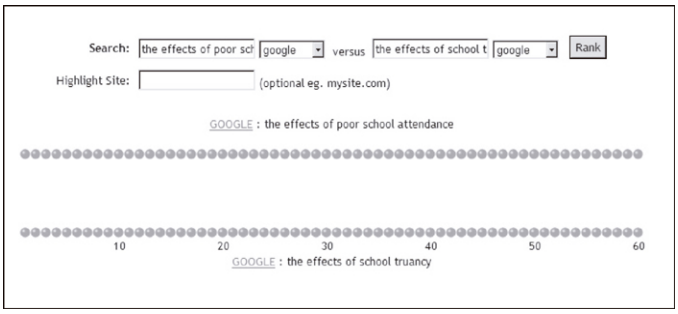
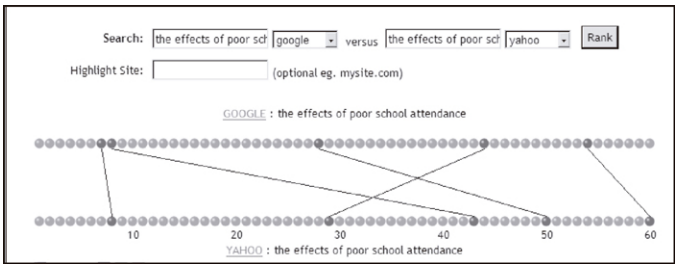


Figure 4.6 Comparison of results for exactly the same search using Google and Yahoo



The fact that there are no lines linking any of the balls on the different rows means that, in the first 60 hits, no common information source was retrieved by both searches. (Since search engines are constantly updating their indexes, you may get different results from those obtained at the time of writing.)

A similar result is obtained if we search for:

looking for information

and:

seeking information

[p. 47 ↓] Different results may also be obtained when *exactly the same search* is performed using *different search engines.* Searching for:

the effects of poor school attendance

in both *Google* and *Yahoo* produces only five items in common – and they are not ranked in the same position. The results are shown in Figure 4.6. As previously noted, lines between the retrieved items indicate that they are identical. The position of each ball in the row reflects its position in the list of hits.

[p. 48 ↓]

Think

Why do you think this happens? How is it that two searches for essentially *the same* information result in completely different hit lists?

To understand how this happens you need to know something about how search engines work. If you are aware of their limitations, you can take steps to compensate for these weaknesses and work around them.

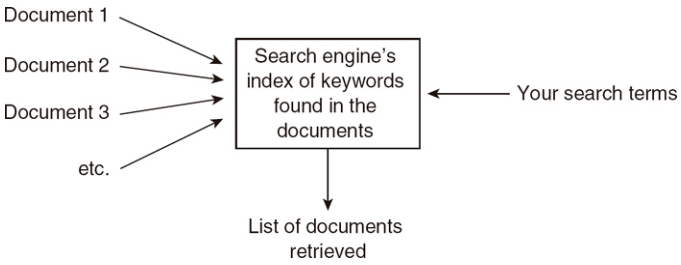
What You Need to Know About Search Engines

Search engines compile huge indexes of the information to which they provide access. These are compiled automatically by computer programs that trawl the web looking for new documents posted on websites.

When they find a new document, they copy words that the document contains and put these into their index, along with a link to the document. When you search, you type words (search terms) into the search box of the search engine, describing what it is you are looking for. These search terms are matched by the search engine with words in its

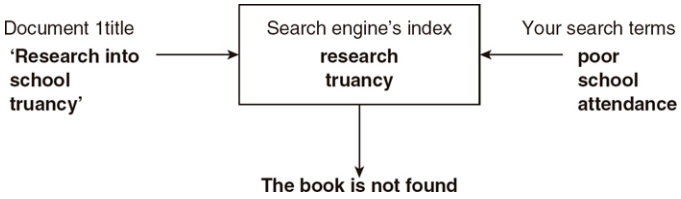
index. Those documents that contain the terms that match your query are the ones that are shown to you. This process is shown in Figure 4.7.

Figure 4.7 Basic search engine indexing and retrieval



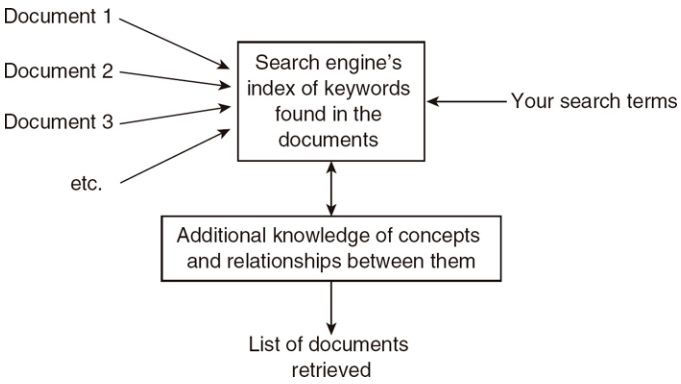
[p. 49 ↓] At the simplest level, what this means is that if a document is indexed by a search engine using the word *school truancy*, and you search for it using the words *poor school attendance*, there will be no match between index entry and search query and the document will not be retrieved (as shown in Figure 4.8).

Figure 4.8 A mismatch between search engine index terms and a user's search terms



As we will see in Chapter 7, search engineers have developed – and are continuing to develop – techniques to lessen this limitation. Their goal is to enable search engines to act more intelligently so that they behave a little more like the lecturer mentioned at the beginning of this chapter, who was able to recommend a book on *school truancy* to the student searching for information on *poor school attendance.* A number of such techniques will be introduced in Chapter 7. The basic idea is shown in Figure 4.9.

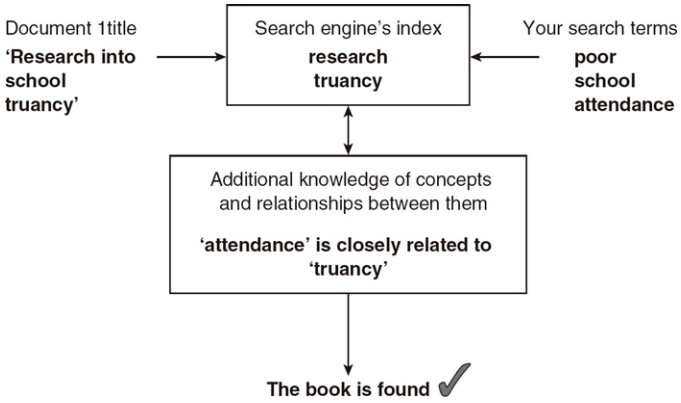
Figure 4.9 Enhanced search engine indexing and retrieval…



If we can provide the search engine with some of the basic common-sense knowledge that humans possess – in this case, the fact that *truancy* and *attendance* are closely related topics, and people searching for one are likely to be interested in the other – it will behave a little more intelligently, as shown in Figure 4.10 overleaf.

[p. 50 ↓]

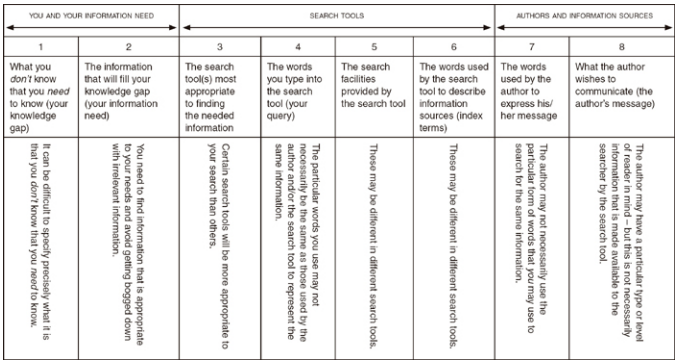
Figure 4.10 … leading to improved search performance



Many search tools offer a range of features designed to try to help the information seeker find the most relevant information for his or her needs. Research is also ongoing to try to build even greater levels of intelligence into search engines – for example, making use of ontologies and semantic web developments.

However, you should be aware that, for the moment at least, the ability of search engines to make intelligent connections is limited. There is still a need to develop your own search skills – for example, to figure out what alternative words and phrases might be used by authors writing about what it is you need to know, and to try different ones as necessary to refine and improve the effectiveness of your search. Figure 4.11 shows some of the obstacles that can stand in the way between you and the information that you need.

Figure 4.11 Obstacles that can prevent you from finding the information you need



Between column 1 (the gap in your knowledge requiring information to fill it) and column 8 (an author's knowledge that can potentially fill your knowledge gap), there are many factors that can intervene to prevent you finding the information you need. However, one way in which you can mitigate some of these problems is to develop effective information-seeking skills. These include being able to:

- identify the type of information most appropriate to answering your essay or research question;
- select the search tools best suited to finding this information;
- devise a search strategy that will make best use of the facilities offered by the search tools in order to maximise the effectiveness of your searching; and
- monitor, evaluate and where necessary improve your searching over time.

These skills will enable you not only to be aware of the limitations and strengths of different search tools, but also to know how to compensate for their limitations and

exploit their strengths to your advantage. The following chapters will explore how you can develop these skills.

[p. 51 ↓]

[p. 52 ↓]

Summary

In this chapter, we defined what is meant by *authoritative information.* This is important since establishing the authority of the ideas you put forward in your work is a necessary feature of academic work. *Peer review* is a mechanism for establishing the authority of an information source, and you should generally find and use peer-reviewed sources.

However, if you are exploring the latest cutting-edge new ideas relating to a topic, these may often be not yet available in peer-reviewed sources. There may be a trade-off between novelty and creativity on the one hand, and consensus and established authority on the other. Unless you are working at an advanced level, you are advised to err on the side of caution and to use peer-reviewed sources wherever possible. Where you do make use of less authoritative sources, you will be expected to justify this use, critically evaluating the sources and providing evidence yourself of why they are valid and appropriate to put forward as evidence in your work.

The chapter went on to explain some of the problems potentially preventing you from finding the information you are looking for. Some of these problems are inherent in the way search tools index the information to which they provide access.

You need to be aware of these limitations so that you can develop your own information skills in order to compensate for them. These skills include being familiar with the range of search tools available to you, and knowing which are most suitable for helping you find the type of information you need. Different search tools also offer a variety of sophisticated techniques for maximising the power and efficiency of your searching, and you need to know what these are and how to use them. These issues will be introduced in Chapters 5 to 8.

http://dx.doi.org/10.4135/9781446287927.n4