# Citation of Sources

CSE417 Communication and Research Skills
CSSE Monash University

## 1 Introduction

This document presents an overview of the principles that should be observed when citing sources, published or otherwise, in both your literature review and your thesis. Additional and supporting information is available in the *Student Guide to Writing Assignments* (Eaves and Hurst, 1998) (hereafter referred to as the Student Guide). You should read the Student Guide as well as this document. It provides a great deal of additional useful information and guidance.

## 2 Citations: What and Why?

The Student Guide says of citations:

> Citations are the references to others' work that you have used. They are placed in the body of your paper. Their purpose is to provide the reader with an intellectual 'audit trail' that can be followed back from your efforts to the sources you used. For this reason, citations and a matching list of references are necessary where a particular published or unpublished source has provided ideas you have used, *even if you have not used a direct quotation*. [emphasis added]

Correctly citing the sources that you have used is a *vital* part of academic writing. Just as important—if not more so—is the correct use of quotation marks " ", for short quotations, or indentation, for long quoted passages (as below), to indicate when text has been taken directly from one of the sources used. As the Student Guide says:

> If you use *more than five words* taken directly and word for word from a published source, and do not indicate that the material is a quotation, you are guilty of **plagiarism**. This is the case even when you cite the source in the body of your text, or in your references, or both. Plagiarism of any extent is severely frowned upon, and where intent to gain an unfair advantage is evident, becomes **cheating**. Examiners are within their rights to award 0 for any work that entails plagiarism, and *cases of cheating can lead to a fail in the subject, or even exclusion.* [emphasis added]

These are not idle threats. The Faculty and the University regard plagiarism and cheating as very serious offences. We want to help you to get it right, but you must take responsibility for your work, and make use of the resources provided and to which you are directed.

### 2.1 Why Are Citations Required?

All academic writing includes and requires citations. Citations show readers where ideas have come from, and make it easy for them to find the original source if they want to know more about it. They are also an indication of the author's scholarship: the appearance of many citations can indicate that the author has read widely and is familiar with the state-of-the-art in his/her field of research.

All research builds on earlier work. Sir Isaac Newton captured this notion beautifully in the sentence “If I have seen further, it is by standing on [the] shoulders of giants” (Newton, 1676). Citing earlier research makes the context of new research clear. You will find that all the academic papers you have located for your literature review cite earlier work for this reason, even if they introduce completely new ideas. Non-academic sources, such as industry ‘white papers’ often do not cite their sources, which means that they are of very limited value.

You cannot make unsubstantiated claims in academic writing. Every claim about the work of others must be supported by a citation. Every original claim, *i.e.* your own work, must be supported by argument, proof or experimental results.

In a literature review, such as the one that you are writing for this unit, new research is not expected. You are expected to read a variety of publications relevant to your topic, to describe their contents, and then to discuss them, comparing the various approaches to problems and discussing their advantages and disadvantages. Since no original research is expected, it is reasonable to expect that almost every paragraph will contain one or more citations. The only likely exceptions to this are the introduction and/or the conclusion, which may consist of general discussion of the problem.

Some students seem to believe that placing many citations in their work is bad, since (they think that) it makes it look as if they have not done anything original themselves. This is *not* the case! The presence of many citations is *good*, since it indicates that you have read widely, and that you understand the importance of acknowledging sources in academic writing.

## 2.2 Paraphrasing and Quotation

Although the presence of many citations is good, it is *extremely* important that you write about the articles that you have read *in your own words*; this is called *paraphrasing* and is normal academic practice (see section 5 of the Student Guide). Whenever you take words *directly* from a source—*i.e.* without rewriting in your own words—you are *quoting* that source. Some quotation in a research paper is fine: sometimes an author seems to have found exactly the right words (as with Newton (1676) above). Quotations *must* be either enclosed in quotation marks or shown as an indented passage. Remember, however, that a paper or thesis consisting entirely or largely of quotations will fail: your supervisor or examiner is interested in *your* work, not simply the words of others pasted together.

# 3 Information Required for Citations

Each citation in the text of your paper must correspond to an entry in the 'References' section at the end. Remember that one of the primary functions of a citation is to allow the reader to find the source that you used. Consequently the entries in the 'References' section must contain enough information for this to be possible. The main questions you must answer are:

- **Who** wrote this?
- **What** is it called?
- **Where** was it published?[1]
- **When** was it published?

Below is a list, taken from Lamport (1994), of possible BIBTEX entry types and the required and optional information that must be supplied when referring to a source of that type. Even if you do not use LATEX and BIBTEX this is an excellent guide to the information that should be provided when citing various kinds of publications.

**article** An article from a journal or magazine.
Required information: author, title, journal, year
Optional information: volume, number, pages, month, note, key, url

**book** A book with an explicit publisher.
Required information: author or editor, title, publisher, year
Optional information: volume, series, address, edition, month, note, key, url

**booklet** A work that is printed and bound, but without a named publisher or sponsoring institution.
Required information: title
Optional information: author, howpublished, address, month, year, note, key, url

**inproceedings** (or **conference**) An article in a conference proceedings.
Required information: author, title, booktitle, year
Optional information: editor, pages, organization, publisher, address, month, note, key, url

**inbook** A part of a book, which may be a chapter (or section or whatever) and/or a range of pages.
Required information: author or editor, title, chapter and/or pages, publisher, year
Optional information: volume, series, address, edition, month, note, key, url

**incollection** A part of a book having its own title.
Required information: author, title, booktitle, year
Optional information: editor, pages, organization, publisher, address, month, note, key, url

**manual** Technical documentation.
Required information: title
Optional information: author, organization, address, edition, month, year, note, key, url

[1] *Not* where did you find it!

**phdthesis** (or **mastersthesis**) A Ph.D. or Master's thesis.
Required information: author, title, school, year
Optional information: address, month, note, key, url

**proceedings** The proceedings of a conference.
Required information: title, year
Optional information: editor, publisher, organization, address, month, note, key, url

**techreport** A report published by a school or other institution, usually numbered within a series.
Required information: author, title, institution, year
Optional information: type, number, address, month, note, key, url

**unpublished** A document having an author and title, but not formally published.
Required information: author, title, note
Optional information: month, year, key, url

**misc** Use this type when nothing else fits.
Required information: howpublished (not really required, but strongly recommended)
Optional information: author, title, month, year, note, key, url

Note that the 'required information' above indicates the *minimum* required. You should supply as much information as you can. The 'References' section at the end of this document gives examples of how an journal article (Borges and Levene, 2000), a conference article (Agrawal and Srikant, 1995), a book (Westphal and Blaxton, 1998) and an on-line industry or government white paper should be cited (Houle et al., 2000).

### 3.1 Citing On-line Sources

Many students find sources for their theses on-line. This is fine, but it is important to be careful when citing sources found on-line, for two reasons:

- often the on-line version is not the primary medium in which the article was published (especially for journal and conference articles). You must find out where the article was actually published and cite those details. Publication in a conference or journal generally implies peer review and thus quality control. Anyone, however, can put up a web page;
- URLs are subject to frequent change. Often documents disappear from the web, or are moved to a new URL. A recent study of articles in five journals found that of more than 1,000 citations including URLs, 33% of the URLs were dead, and only 56% of the live URLs still pointed to someting relevant to the citation (Carlson, 2005).

The second point above means that it is **not acceptable** simply to give a URL for an on-line document. First you must try to find out if it has been published elsewhere (almost all academic articles have been), and if you cannot find any such information, you must give at least the author(s) and title of the document, as well as the date on which it was downloaded. See the Student Guide for directions on citing articles with no apparent author.

### 3.1.1 The Digital Object Identifier System

In recent years many major academic publishers (*e.g.* Elsevier, Kluwer) have started to use the **Digital Object Identifier (DOI)** system. A DOI provides a unique identifer for a digital object, such as a journal or conference article. DOIs are guarranteed not to change. If a DOI exists for an article it should be used in preference to a URL, *e.g.* `doi:10.1000/123`, and/or the URL should be given in terms of a DOI, *e.g.* `http://dx.doi.org/10.1000/123`. Read more at `http://doi.org/`.

# 4 Tools to Help Manage Citations

Whilst it is reasonably easy to manage citations and references for a small essay or paper manually, it becomes much more difficult when writing large papers, theses or books. Authors of such works use tools to help then with this task.

The document preparation system used by many scientists, particularly in computer science, physics and mathematics, is LaTeX. It has a bibliography management system called BibTeX, and is also excellent for managing large documents and typesetting mathematical formulae. LaTeX is available on most, if not all, UNIX systems (including Linux). It is also available for MS Windows. There is further information available on-line at: `http://www.csse.monash.edu.au/software/latex/`, including the `cssethesis` document class, and example files.

If you insist on using MS Windows software, there is a tool for MS Word and Word Perfect called *EndNote*. It is available free to Monash staff and students. Details on how to obtain and use EndNote are available on-line at:
`http://www.lib.monash.edu.au/vl/endnote/endncon.htm`

## 4.1 The natbib Citation and Bibliography Package

The natbib package provides extended citation capabilities (Daly, 1999). In particular it supports the Harvard "author, date" citation style that is preferred for theses, when used with an appropriate bibliography style, such as `dcu`. It replaces the standard LaTeX `\cite` command with two basic forms of citation command: `\citep` and `\citet`, as well as providing several other very useful ones.

The `\citep` command is best used when placing a citation at the end of sentence or phrase (as above). In the natbib documentation, this is referred to as a *parenthetical citation*.[2]

When you want to refer to the authors of a particular work, typically at the start of a sentence, a parenthetical citation is not appropriate. This is particularly so if you are using a numerical or symbolic citation style. You should *not* start a sentence with

> [2] says that this is most certainly ...

In such situations you really need to give the authors' names. The `\citet` command produces *textual citations*, which allows you to produce things like:

[2] For ease of conversion from exisiting LaTeX documents, you might find it useful to place `\renewcommand{\cite}{\citep}` in the preamble of the document, since any existing standard `\cite` commands should almost certainly be treated as `\citep`.

> Ade (1983) describes a means by which textures may be characterized . . . another approach is given in De Bonet and Viola (1998).
>
> Atick et al. (1996) note that humans have little or no difficulty in perceiving shape, yet find it extremely difficult to *describe* what they perceive.

Note that an abbreviated version of the authors' names has been produced in the third example above. It is often desirable to have the full list of authors' names given when a work is first cited, and an abbreviated list thereafter. This can be achieved by passing the `longnamesfirst` option to natbib when the package is used. This will produce an initial citation like:

> Atick, Griffin and Redlich (1996) note that humans have little or no difficulty in perceiving shape, yet find it extremely difficult to *describe* what they perceive.

Both the `\citep` and `\citet` can take two optional arguments. If just one is provided, its text will appear as a "post-note" after the citation details. If two arguments are provided, the first defines a pre-note, and the second a post-note. Here is an example:

> `\citep[Ch.~3]{AaK1989}` . . . (Aarts and Korst, 1989, Ch. 3)
> `\citep[see][Ch.~3]{AaK1989}` . . . (see Aarts and Korst, 1989, Ch. 3)

If you want to use a numerical citation style, *e.g.* [2] rather than the Harvard (author, year), but have `\citet` and `\citep` working, pass the option `numbers` to the natbib package:

```
\usepackage[numbers]{natbib}
```

and use either `plainnat` or `unsrtnat` as the bibliography style.
These examples only scratch the surface of what the natbib package can do. To discover the full power of the package, see the documentation at CTAN (Daly, 1999). You probably already have it on your system. Try `locate natbib.dvi` at the command line.

# 5 Common BIBTEX Mistakes

## 5.1 Author lists

BIBTEX requires that the names of authors in the "author" field be separated by the word `and`. They must *not* be comma-separated. Two name formats are supported. You may use

```
author = {David McG. Squire and Fred Quentin Nurke and Karen {Sparck Jones}},
```

or

```
author = {Squire, David McG. and Nurke, Fred Quentin and Sparck Jones, Karen},
```

These result in citations (Squire, Nurke and Sparck Jones, 2002a) and (Squire, Nurke and Sparck Jones, 2002b) respectively. Note the use of the parentheses `{}` around the two-part surname `Sparck Jones`. They are necessary when using the first format, but not the second (since everything before the comma is treated as the surname). If they are omitted from the first, the citation (Squire, Nurke and Jones, 2002) results.

### 5.2 Capital letters in the `title` field

Many BIBTEX bibliography styles automatically make all but the first letter of the `title` field lowercase. This is a problem if you need to use a proper noun, or an abbreviation, in a title. Letters for which the case should be preserved must be enclosed in parentheses. The BIBTEX entry

```
@misc{SNS2002a,
author = {Squire, David McG. and Nurke, Fred Quentin and Sparck Jones, Karen},
title = {Something about {XML}},
year = {2002},
}
```

produces the citation (Squire, Nurke and Sparck Jones, 2002b), whereas

```
@misc{SNS2002b,
author = {Squire, David McG. and Nurke, Fred Quentin and Karen Sparck Jones},
title = {Something about XML},
year = {2002},
}
```

does not retain the correct uppercase for XML, as can be seen in the References section for (Squire, Nurke and Jones, 2002).

## References

Aarts, E. and Korst, J. (1989). *Simulated annealing and Boltzmann machines*, John Wiley and Sons, New York.

Ade, F. (1983). Characterisation of textures by "eigenfilters", *Signal Processing* **5**: 451–457.

Agrawal, R. and Srikant, R. (1995). Mining sequential patterns, *in* P. S. Yu and A. L. P. Chen (eds), *Proceedings of the 11th International Conference on Data Engineering*, Taipei, Taiwan, pp. 3–14.

Atick, J. J., Griffin, P. A. and Redlich, A. N. (1996). The vocabulary of shape: principal shapes for probing perception and neural response, *Neural Computation* **7**(1): 1–5.

Borges, J. and Levene, M. (2000). A fine grained heuristic to capture web navigation patterns, *ACM SIGKDD Explorations* **2**(1): 40–50.

Carlson, S. (2005). Scholars note 'decay' of citations to online references, *The Chronicle of Higher Education* **51**(28): A.30.

Daly, P. W. (1999). Natural sciences citations and references (author–year and numerical schemes), LATEX2e package documentation. (last accessed February 7, 2004).
**URL:** *http://www.ctan.org/tex-archive/macros/latex/contrib/natbib/*

De Bonet, J. S. and Viola, P. (1998). Texture recognition using a non-parametric multi-scale statistical model, *Proceedings of the 1998 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'98)*, Santa Barbara, California, USA.
**URL:** *http://www.ai.mit.edu/~jsd/research/publications/1998/DeBonet-CVPR98.pdf*

Eaves, D. and Hurst, J. (1998). Standards for the presentation of written assignments, School of Computer Science and Software Engineering, Faculty of Information Technology, Monash University, Australia; On-line resource. (last accessed June 8, 2005).
**URL:** *http://www.csse.monash.edu.au/~ajh/teaching/resources/studentguide.pdf*

Houle, J. L., Cadigan, W., Henry, S., Pinnamaneni, A. and Lundahl, S. (2000). Database mining in the human genome initiative, Bio-databases.com white paper, Amita Corporation, 1420 Blair Place, Suite 500, Ottawa, Ontario, Canada, K1J 9L8.
**URL:** *http://www.biodatabases.com/whitepaper01.html*

Lamport, L. (1994). *LaTeX: a document preparation system*, second edn, Addison-Wesley Publishing Company, Inc., Reading, MA, USA.

Newton, I. (1676). Letter to Hooke.

Squire, D. M., Nurke, F. Q. and Jones, K. S. (2002). Something about xml.

Squire, D. M., Nurke, F. Q. and Sparck Jones, K. (2002a). Something about XML.

Squire, D. M., Nurke, F. Q. and Sparck Jones, K. (2002b). Something about XML.

Westphal, C. R. and Blaxton, T. A. (1998). *Data Mining Solutions: Methods and Tools for Solving Real-World Problems*, John Wiley & Sons, New York, USA.