

International Area Studies Review

<http://ias.sagepub.com/>

Active citation through hyperlinks: The retarded replication revolution

Stein Tønnesson

International Area Studies Review 2012 15: 83

DOI: 10.1177/2233865912437692

The online version of this article can be found at:

<http://ias.sagepub.com/content/15/1/83>

Published by:

Center for International Area Studies, Hankuk University of Foreign Studies



Peace Research Institute Oslo



and

<http://www.sagepublications.com>

Additional services and information for *International Area Studies Review* can be found at:

Email Alerts: <http://ias.sagepub.com/cgi/alerts>

Subscriptions: <http://ias.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations: <http://ias.sagepub.com/content/15/1/83.refs.html>

>> [Version of Record](#) - Mar 29, 2012

[What is This?](#)

Active citation through hyperlinks: The retarded replication revolution

Stein Tønnesson

Peace Research Institute Oslo (PRIO), Norway, and Department of Peace and Conflict Research,
Uppsala University, Sweden

Abstract

The hyperlink revolution in academic referencing is long overdue. Why is the potential offered by the internet for providing hyperlinks to sources so little utilized by scholars? The article discusses the most likely reasons and suggests that the revolution will happen only when it is promoted from above through mandatory hyperlinks in term papers, dissertations, journal articles and scholarly monographs. This will make it possible for readers to replicate, test and evaluate not just statistical analyses of quantifiable data, but any argument based on qualitative or quantitative analysis of any kind of source, be it an object, text, speech act, image, or film.

Keywords

References, hyperlinks, replication, active citation, qualitative research

Hyperlinks have a potential to revolutionize scholarly reference systems. It has for a long time already been possible to provide direct links to sources of virtually any kind. Thus it becomes much easier for others to check or replicate the arguments and calculations we use when drawing our conclusions. Yet our traditional reference systems have not been reformed to fit the new opportunities. Most scholars shy away from using hyperlinks, and students are not systematically encouraged to use them. Paradoxically, the abundance of electronically available material and the software we use to access and organize it sometimes seem to enhance vagueness instead of precision in citations. This article discusses why the revolution in academic referencing is lagging, and suggests what should be done.

Elaborate reference systems are a distinctive mark of academic scholarship, both qualitative and quantitative. A scholarly book or article presenting findings based on empirical research must

Corresponding author:

Stein Tønnesson, Peace Research Institute Oslo (PRIO), Box 9229 Grønland, Oslo, N-0134, Norway.
Email: stein@prio.no

provide readers with references to where the evidence can be found. An article based on statistical analysis is expected to provide readers not just with a description of its method but also direct access to the data that went into the analysis and the mathematical equations or models used. Only thus can research be replicated.¹ The same goes for philosophical and other theorizing: An article presenting a new idea or theory inspired by pre-existing ideas should reveal through precise references where the inspiration came from.

Scholarly disciplines have different reference systems. Social scientists mostly prefer the Harvard format with short references in parentheses inside the text: author name, year of publication, page numbers related to a list of references with all the most relevant bibliographical data (e.g., 'Wang Jieshi 2010: 44'). Anthropologists and sociologists use various ways of referring to interviews or surveys (e.g., 'Interview with village headman, Pattani province, 18 Dec 2012'). Historians mostly prefer to use footnotes since they are more suitable for references to archival documents (e.g., 'Churchill to Roosevelt No. 365B, 1 Apr 1945, CAB3615, The National Archives, London'). The system used by *International Area Studies Review* is ideal for a multi-disciplinary journal since it allows both Harvard style citations in parentheses and numbered notes at the end of its articles, and then a list of references.²

There are two main features that all reference systems have in common: (a) their primary purpose is to make it as easy as possible for someone else to access the same source and replicate the same calculation or logical argument; (b) this purpose is radically enhanced by the provision of hyperlinks directly to sources.

There is virtually no limit to the kind of sources that can be accessed through web-based hyperlinks. Wikipedia is a primary example. Links can point towards a spreadsheet with a dataset, a table, an equation, a published document, notes taken during an interview, a recording of an interview, a film, radio broadcast, photograph, or a copy of an archival document. Past events as such cannot of course be replicated the way you do with a scientific experiment, but this was also the case when scholars had to rely on paper-based citation systems. All that has ever been cited on paper can now be accessed directly through a web-based hyperlink if only someone scans it and posts it on the web.

Ten years have passed since Wikipedia was launched on 15 January 2000.³ Yet it is safe to say that most academic publications, at least in the social sciences and humanities, do not include hyperlinks to sources. A huge majority of scholarly books and articles are still published without any accompanying website providing hyperlinked access to the sources. Readers are thus in no better position to test scholarly findings now than they were in the age before the internet. This is mainly true of qualitative research. Scholars using quantitative methods are somewhat more advanced in using new replication technology than those of us who rely mostly on qualitative methods.

The missing revolution from below

The historically oriented political scientist Andrew Moravcsik has contributed to setting off a self-critical reflection among his colleagues on the crisis of qualitative case studies, a crisis he attributes to a 'failure to impose firm standards of replicability' (Moravcsik, 2010: 29). His proposed recipe is to require of political scientists that they back up any critical and contested substantive empirical point in a scholarly case study by 'a precise and annotated citation to one or more presumptively primary sources' (Moravcsik, 2010: 31). This is a classic requirement within the discipline of history. History is not in the same kind of 'self-imposed crisis' as the one Moravcsik describes for

political science, but most historians suffer from a lack of attention to the opportunities offered by the internet, and from insufficient interaction with both quantitatively and qualitatively oriented social sciences.

A few historians have made modest attempts to publish books or articles in combination with a web-based publication, but not to the extent one would expect. A highly beneficial replication revolution for which there is no technological hindrance simply does not seem to be happening. I am not now thinking of e-books. They are actually a problem since they lack page numbers and cannot be directly hyperlinked. The *Chicago Manual* thus allows references to sections instead of page numbers. This is a huge paradox since it means that checking sources becomes more instead of less difficult than before.⁴ What I suggest as a new requirement is that every scholarly publication must be adjoined to a set of webpages providing links to its sources. Some scholars have done this on their own initiative, most often when being asked by a publisher to cut down their manuscripts.

Let me take an international historian (myself) as an example. The text in Tønnesson (2010) included numerous quotes from documents collected over a period of thirty years in French, British, US and Vietnamese archives. I translated many of these quotations from French and some of them from Vietnamese (with help from Vietnamese colleagues) but wanted to give readers access to the quotes also in their original language. Sometimes I also wanted readers to get a chance to read much longer excerpts from certain documents than the ones I could include in my printed narrative. Hence I chose to dump all this extra material into a massive number of footnotes. Once the publisher had read my manuscript, it came as no surprise when the acceptance letter underlined the necessity of cutting these footnotes down to a minimum. I suggested then that the publisher allow me to include a reference in the book's preface to my own personal website (www.cliostein.com), where I would publish the full, or extended footnotes. Readers could then decide either to read these footnotes on a monitor or keep a printout next to the book while reading. The University of California Press accepted my proposal and even allowed me to include more footnotes in the printed volume than was necessary, so their numbering would correspond to the numbering of the extended footnotes on the web.

Once we had agreed on this solution I realized that the website could be used also for other purposes. One was to publish a list of errata. Although the mistakes discovered were corrected by the publisher in the paperback that came out a year after the hardback version, I felt that buyers of the hardback version would benefit from such a list. A second was to publish links to the reviews of the book that were published in various journals, and notably to an electronic roundtable organized by H-Diplo, the H-NET discussion list dedicated to the study of diplomatic and international history. Third was to publish a blog with my own reflections on the book's reception, notably on the question of the inevitability of the Indochina Wars.

Then the fourth and biggest possibility dawned upon me: provide the extended footnotes with hyperlinks to the full archival documents. This would give any reader of the book a chance to look up any footnote on the web and then access the full source through a hyperlink. When I realized how useful this would be, I started wondering why so few scholars had yet availed themselves of this opportunity. Then I started wondering if the best solution might be to simply omit all reference notes from the printed version of a scholarly monograph, and publish the references only on the web. Then instead of having numbered notes one could refer to a page number in the book and to a sequence of words occurring on that page. This kind of reference system has long been used by Random House in books aimed at a wide audience.⁵ An example of what it looks like can be found in Harries and Harries (1991: 514):

CHAPTER 13: COUNTERATTACK

- 135 Munitions Bureau and Council . . . abolished: WO 106/5479, Feb. 25, 1925.
135. 'We must substitute': A. Iriye, 'The Failure of Economic Expansionism 1918–31,' in B. Silberman and H. Haratoonian, eds., *Japan in Crisis: Essays on Taisho Democracy* (Princeton, N.J.: Princeton University Press, 1974). p. 245.
135. Baron Oi: WO 106/5479, June 29, 1925.⁶
136. For Germany in 1923, see G.A. Craig, *op. cit.*, pp. 450, 461.
136. The Kaikosha-Kiji articles were discussed in our interview with Toshio Morimatsu, May 1989.

While this system cannot really compete with the numbered system when the notes are published only on paper, it is ideally suited to a new win-win solution, where publishers can print books without any notes at all (although they must have page numbers), and instead let every book be accompanied by an expandable and updatable reference document on the web.

Since I believe that printed monographs still have a bright future – and that publishers will continue to protect e-books from access through hyperlinks – the combination of a printed volume and an author- or university-controlled website is more likely to be a successful solution than full electronic integration on the Wikipedia model.

So far I have not found the time or resources needed to provide hyperlinks in a systematic fashion from my own extended footnotes to my personal collection of archival documents. I have only linked a few of my footnotes to copies of documents for which readers have shown particular interest. When I have received an email with questions concerning a specific reference, I have posted the original document on the web and inserted a hyperlink to it in the web-based extended footnote document instead of just sending it to the person who wrote me the email.

My private collection of archival documents does not of course contain any originals. They are in the official archives. My footnotes indicate the number and/or title of the folder, box and file where the original documents are located. My collection consists of (a) notes written by hand in the early 1980s, (b) notes taken on a personal computer from 1984 onward, (c) photocopies of documents, which can be scanned, and (d) in a few instances digital photographs (the main method used in archival research today). My personal archive is fairly typical. Most historians of my age have similar collections of source material. It is a colossal waste that such collections are exploited only by one researcher for the duration of his/her lifetime, with many documents being thrown out as garbage or stored in an inaccessible place when their owner passes away.⁷ The same is the case for qualitative research in the social sciences: ' . . . most qualitative data generated by American social scientists are used only once. Scholars wishing to retrace the steps taken by others typically must repeat the research' (Elman et al., 2010: 23).

When my book was published, I launched the website publication of the extended footnotes at the same time, and circulated a letter to all of my email contacts to inform them of the accompanying website. In that letter, dated 2 October 2009, I was naïvely optimistic:

Inspired by the *Journal of Peace Research's* policy on replication data I've come to think that it makes sense to publish references to source material on the net also for historical monographs. . . . In a not too distant future, I think most historical monographs and articles in history journals will be published with footnotes that include hyperlinks to copies of the original documents. This will make it much easier to check how our colleagues use and interpret their sources.

I got a number of positive replies, and many respondents shared my belief that a revolution in referencing was on its way. Some thought it might improve the ‘honesty’ of researchers. I learned that the international historian Marc Trachtenberg had long been doing something similar. Already in 1998, when his publisher demanded that he cut his book *A Constructed Peace* down to a maximum of 235,000 words, he published a set of appendices on the web, and has since published appendices to his later works as well on his personal website.⁸ However, there were replies to my email suggesting I was overly optimistic. They thought most scholars would be reluctant to assume the extra burden of making their sources available on the web.

The most interesting of all the responses to my 2 October 2009 circulation letter came from Andrew Moravcsik of Princeton University. When he received my letter, he was just about to complete his article ‘Active Citation: A Precondition for Replicable Qualitative Research’. It put forward almost the same suggestion as I did, namely to set a new standard for referencing in qualitative/historical research by utilizing the opportunity to publish ‘virtual versions of articles’ on the web, including text from all primary sources – or the sources themselves. This is what Moravcsik calls ‘active citation,’ which he defines as ‘rigorous, annotated (presumptively) primary-source citations hyperlinked to the sources themselves’ (Moravcsik, 2010: 31). ‘The idea,’ Moravcsik explained in his letter to me, ‘is that our work should be as replicable as that of our quantitative colleagues, thus stimulating criticism, debate, emulation, extension, and – I believe ultimately, at least within the social sciences – more respect in the field.’ Moravcsik said he would like to believe in my prediction that most historical monographs and articles were soon going to be published with footnotes including hyperlinks to copies of the original documents. Yet he saw ‘little evidence that it is becoming true yet. Even in political science it’s almost unheard of; only in the experimental social sciences and, of course, natural sciences, do you start to encounter it.’ Then he asked me if I could provide him with an example of a historical journal moving in the preferred direction.⁹

I could not.

Why is the revolution lagging?

Why is it that scholars have not jumped into the hyperlinked world in the way of the enthusiasts behind the Wikipedia miracle? From an academic perspective, the advantage seems so obvious. It will dramatically increase scholarly accountability, reduce the need for double work, enhance the publication of document collections on the web, reduce the opportunity for academic fraud, and perhaps help cure some of the illnesses that plague scholarly referencing today, particularly among sociologists and political scientists. The practice of referring to a full article or monograph instead of a precise page number has become quite widespread.¹⁰ The rationale used to defend this lazy habit is that you sometimes need to refer to a book’s general idea or argument rather than anything in particular. I do not wish to claim that this may not sometimes be justified, but it is my impression that the argument often serves as an excuse for referring to works one has not actually read. The real purpose sometimes seems to simply demonstrate awareness that a work exists. Such malpractice may of course continue to thrive in a world of hyperlinked references, but will then be easier to expose. I concur with Moravcsik’s complaint that ‘Current practice in some areas of political science permits citations to be imprecise, vague, and secondary, rather than precise, annotated, and primary.’ The ease with which one can add more references through electronic access to published articles has loosened the attachment between text and source: ‘As scientific citations spread and word limits tighten,’ adds Moravcsik, ‘references increasingly lack any precise annotation describing how the evidence supports the textual claim’ (Moravcsik, 2010: 30).

A first reason why the replication revolution has not yet really happened could be simple inertia. Scholars tend to maintain their old habits and teach them to their students. Such conservatism is more prevalent in some scholarly disciplines than others. It is interesting to notice that the discipline of history is more conservative than law. There has been a real revolution in access to legal sources. A driving force here has been commercial. Legal databases offer huge commercial opportunities. The opportunity to charge money for access to historical databases is not on the same level. Yet the number of databases with primary historical source material is also now rapidly increasing.¹¹

A second reason is that the commercial interest of publishers does not extend to creating viable referencing systems. In the world of book publishing these systems are often seen as a hindrance to commercial success. Book publishers are therefore unlikely to become a part of the solution, although they may benefit from a system where printed volumes are published either without or with very short references while the authors themselves or their universities take responsibility for maintaining web-based hyperlinked reference documents.

A third reason why the revolution has not been happening is the extra work a scholar needs to undertake to make the source material available. After getting through the tough work of writing up a publishable article it may be tempting to refrain from spending yet more time on providing access to the sources. Many scholars have not even learned the simple technique of creating a hyperlink.

A fourth reason is that career-wise there is little to gain from doing more than your colleagues to open yourself to their scrutiny. Some may even prefer to keep their source material to themselves since they hope to publish more work in the future on the basis of the same material. It may thus seem preferable to keep it away from rival scholars with different ideas.

Fifth, in some cases open access is hindered by ethical considerations. Archival documents may contain sensitive information that is not suitable for being published directly on the web. In order to get access to a particular archival file, scholars may have to pledge not to reproduce the documents in full, but use them only as a basis for their scholarly analysis. In such cases there may be a need for a formal permission from the archive before a scanned document can be reproduced on the web. The ethical problems may be even more prevalent for structured or semi-structured interviews, oral history, and notes taken during anthropological participatory observation. References to interviews are often anonymized. If colleagues and other readers are to have access to the recording of an interview or the notes taken by a researcher during an interview or fieldwork, then the explicit permission of the interviewee may be needed. A possibility might be for scholars to write a double set of notes, one anonymized with a view to web publication, and another to be kept in each scholar's confidential file.¹² It is easily understandable if such considerations prevent scholars relying on oral interviews from entering the world of hyperlinked referencing. But it is hardly acceptable.

A sixth reason may be concern for the stability of the web. A printed book in a library may survive for centuries or even millennia. Web-based sources come and go with changing software and failures in institutional maintenance. Although understandable, this concern is no longer a valid argument for sticking to paper. It means only that we must ensure long-term preservation of electronic archives. There are systems available that will protect the longevity of web-based material. They must be used by all responsible institutions. What is clear is that we cannot just rely on authors maintaining their own personal websites in the way of Trachtenberg and Tønnesson.

What should be done?

Let me first emphasize that the revolution is already happening in the area of cross-referencing between scholarly journals. EndNote, Reference Manager, Zotero, RefWorks and other software

have become standard tools among students and young scholars; once they are used in conjunction with institutional electronic subscriptions the students are automatically inserted in the hyperlinked academic world. This may further enhance the standing of the electronic journal as the primary channel for academic publishing. However, it also entails a risk that students spend their time reproducing or combining recently published academic work instead of collecting new empirical evidence, reading the classics or developing ideas of their own. In order to prevent such skewed practices it is urgent to make it both easy and mandatory to insert hyperlinks to primary sources as well as to books and articles published long ago.

It has become my conviction that a sustained institutional or even political effort is needed to bring academia into the hyperlinked world. There is of course already some movement in this direction. The University of Essex in the United Kingdom has been in the forefront of creating a repository for qualitative data; it began this work even before the archives could be digitized, and has served as a model for other European institutions (Elman et al., 2010: 26). Universities and other permanent institutions must take responsibility for safe and accessible data archiving. But this is not enough. Universities must also require of their students that they include hyperlinks to sources in their term papers, essays and dissertations, and teach them how to do it properly. Guidelines for academic hiring and tenure must include checking that submitted articles and monographs from applicants include hyperlinks to sources or are accompanied by a web-based source publication. Universities and research institutes should develop clear standards for how to credit systematic data archiving in its systems of promotion, and adapt these standards to the needs of each scholarly discipline.

The replication revolution is long overdue. Both sticks and carrots are needed to make it happen.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Notes

- 1 Some journals, such as the *Journal of Peace Research*, require this of authors using quantitative methods.
- 2 The references in this article are hyperlinked to the journal articles cited.
- 3 http://en.wikipedia.org/wiki/History_of_Wikipedia (accessed 21 December 2011).
- 4 This was discussed on the discussion list H-Diplo in January 2012.
- 5 A recent example is Morgan, 2010: 653–692.
- 6 WO is an acronym for the archives of the United States War Office.
- 7 Davenport (2010) provides a vivid description of how sensitive and diverse the personal archive of even a relatively young researcher can be, and how such archives reflect the various phases of a research career and intellectual trajectory. It will continue to be important for scholars to deposit in protected archives the documents that they cannot for various reasons make public in their lifetime.
- 8 <http://www.sscnet.ucla.edu/polisci/faculty/trachtenberg/> (accessed 12 January 2012).
- 9 Moravcsik to Tønnesson, 2 Oct 2009. Tønnesson's email archive, PRIO.
- 10 As Nygaard (2008: 177) says in her eminent guide to academic writing, 'Regardless of the format [Harvard, Vancouver, Chicago], a page number is almost always required.'
- 11 Griffith (2000) is a wonderful example of how hyperlinks inside a text can provide direct access to archival databases, electronic journals, gateway portals, media, and museums.
- 12 Elman et al. (2010: 25) mention as a problem that ethical concerns may lead to 'storage of partial or incomplete accounts and data sets.'

References

- Davenport C (2010) Data from the dark side: Notes on archiving political conflict and violence. *PS: Political Science & Politics* 43(January): 37–41.
- Elman C, Kapiszewski D and Vinuela L (2010) Qualitative data archiving: Rewards and challenges. *PS: Political Science & Politics* 43(January): 23–27.
- Griffith R (2000) Un-tangling the web of cold war studies; or, How one historian stopped worrying and learned to love the Internet. *Journal of MultiMedia History* 3. Available at: <http://www.albany.edu/jmmh/vol3/untangling/untangling.html> (accessed 22 December 2011).
- Harries M and Harries S (1991) *Soldiers of the Sun: The Rise and Fall of the Imperial Japanese Army*. New York: Random House.
- Moravcsik A (2010) Active citation: A precondition for replicable qualitative research. *PS: Political Science & Politics* 43(January): 29–35.
- Morgan T (2010) *Valley of Death: The Tragedy at Dien Bien Phu that led America into the Vietnam War*. New York: Random House.
- Nygaard L (2008) *Writing for Scholars: A Practical Guide to Making Sense and Being Heard*. Oslo: Universitetsforlaget.
- Tønnesson S (2010) *Vietnam 1946: How the War Began*. Berkeley, CA: University of California Press.