

Consider the Source: Predictors of Online Citation Permanence in Communication
Journals

Daniela V. Dimitrova and Michael Bugeja

***Acknowledgement: The authors would like to thank Hye-Hyun Hong, a master's student in the Greenlee School of Journalism and Communication, Iowa State University, for her assistance on this project.**

Abstract

This study focuses on six leading communication journals and their use of online citations in articles published between 2000 and 2003. The study uses content analysis to explore if there is a relationship between citation characteristics and their stability. The findings show that online citations in the .gov and .org domain are more likely to remain accessible over time. Year of publication and URL level also emerged as significant predictors of online citation permanence. More than 37 percent of the online citations have disappeared from the original source over a four-year period (2000–2003). The implications of these findings are discussed in the context of reliability and replicability of scholarship.

Introduction

Scholars from all academic fields increasingly rely on online sources in their research. Although a number of studies have focused on the use of online bibliographies by librarians,¹ few studies have examined the reliability of online references in the field of communication. This study focuses on refereed research articles published between 2000 and 2003 in the leading communication journals and explores whether there is a relationship between online citation characteristics and their stability. It examines if some online citations are more stable than others, depending on the URL source. First, we review the purpose of the citation, next we synthesize prior research regarding online citations, explain the methodology of the study, and, finally, summarize the results and their implications.

The Goal of the Footnote

The goal of rhetoric is to persuade.² The goal of the footnote is to prove.³ The history of modern footnotes (rather than gloss or marginalia) may date back to Erasmus' *On Copia* (1512) in his useful division of rhetoric into *expression of words* and into *subject matter or ideas* with the latter setting the stage for Francis Bacon (1561–1626) and then John Locke (1632–1704).⁴ Both Bacon and Locke advanced their ideas during the new era of the printing press, which fixed the text and/or the translation of classical works (including those of Aristotle, Cicero, Quintilian, among others) upon which knowledge had been handed down, primarily on standards associated with deduction and expression of words. Indeed, in book one of Aristotle's *Rhetoric*, he asserts that the discipline belongs "to no science" and hence, must be "the function of an art" whose key concepts can be analyzed

methodically.⁵ Thus, Bacon and Locke could scrutinize the source of knowledge and thereby question the value of expression in the search for verifiable truth. Bacon, in leading the new scientific movement of the 17th and early 18th century, posits that meticulous observation and induction would eventually reveal verifiable truth.⁶ Locke picks up on this, questioning the tropes and passions of expression—which he calls "perfect cheats"—that insinuate bad ideas in the search for demonstrable truth.⁷ The works of Bacon and Locke prompted others during the Enlightenment to "reconsider the source," along with the methods of confirming truths in the physical world.⁸

Before footnotes could be used to help elucidate such truths, however, the language had to be fixed not only through printed works but also through dictionaries. By the end of the 17th century, that had occurred in France and Italy but not yet in England (and hence, America) until Samuel Johnson's *Dictionary* (1755). Once both the text and language were fixed via the printing press and dictionaries, Leopold von Ranke (1795–1886) could summon the legacy of Erasmus, Bacon, and Locke to advance historical objectivity and scientific history, both of which have helped to define scholarship and research methodology from primary source to peer review. In its investigations of subjects and ideas, scientific history relies on primary rather than secondary sources.⁹ Thus, the erosion of Internet footnotes—the phenomenon of inaccessible online footnotes—undermines the standards of scholarship and the methods of research, primarily because it destabilizes fixed language and original source. **This reverses the modern emphasis on the verifiable truths of subject matter or ideas, without which we empower unsubstantiated claims.** Today, projects such as Google Scholar and Google Print try to put online and digitize all previous published scholarly work, despite

these and other serious concerns.¹⁰ While more and more information is made available on the Internet, researchers are beginning to notice that digital information is ephemeral and unstable, as the next section demonstrates.

The Half-Life Phenomenon

Scholars have documented the increasing use of online citations by researchers across academic disciplines.¹¹ Mary Rumsey, for example, found in a random sample of law citation reviews that articles citing the Web increased from just 130 in 1995 to 5,462 in 2000. John Markwell and David Brooks documented the problems of "link rot" in the field of biochemistry.¹² Jeffrey Kushkowski found that the persistence of online citations in both print and electronic theses, regardless of academic field, was poor.¹³ Philip Davis and Suzanne Cohen noted that student use of Web references jumped significantly from 1996 to 1999, from 9 percent to 21 percent, respectively.¹⁴ At the same time, scholars have also observed that online citations using URLs as links tend to disappear, and, in some cases, more than half of them fail over time.¹⁵ Different terms are used to describe this phenomenon: some call it link rot others persistency or persistence; availability of online sources or half-life. We use the term half-life to describe this impermanent nature of online citations because it allows us to measure the decay rate over time. Half-life is defined as the time interval during which half of the online citations in a journal disappear.¹⁶

Whereas the half-life varies across disciplines, it has been identified as a pressing issue in many fields, including library and information sciences, legal research, medical

studies, physics, biochemistry, and journalism and mass communication. It has infected undergraduate student papers as well as graduate student theses.¹⁷ Researchers in library and information sciences, however, were the first to show that accessibility of Internet references is a serious problem. Paul Jenkins observed that some fields such as the humanities and social sciences may be more prone to using Web sources than others, which necessitates a deeper and more critical look into humanities and social science journals and scholarship.¹⁸

Predictors of Online Citation Permanence

While many library studies have documented the problem of link rot, few have identified what factors affect the stability of online information. We identify four factors as possible predictors of online citation permanence and test their effects on URL longevity.

Determining some objective characteristics of the online source may be helpful as a way to recognize more reliable online citations from less reliable ones.

Year of publication

In one of the first studies to examine the use of Internet citations by journalism and communication researchers, we identified that the year of publication was significantly related to the citation stability, specifically that citations from most recently published articles were more likely to remain active compared with those cited in older articles.¹⁹ Similar findings were reported by Rumsey in the area of law citation reviews.²⁰ Finally,

David Tyler and Beth McNeil also found that the number of failing online citations increased over time.²¹ Thus, we propose the following:

Hypothesis 1: Online citation permanence will decrease as online citation age increases.

Retrieval dates

Style guides of today do not provide consistent requirements regarding online citations' retrieval dates. Perhaps that is one of the reasons why we found that the majority of the online citations in our study did not include retrieval dates. We also noted that e-citations providing retrieval dates seemed more stable.²² Rumsey concurs that, while there are no easy solutions to the problem of vanishing online sources, it is recommended that authors and editors provide as much information about the citation as they can in parenthetical form to improve the chance of locating the citation in the future.²³

The *Bluebook* is the bible for law researchers. Its stipulations for online sources recommend, first, that scholars use traditional printed sources, if possible, as the preferred source. If that is not possible, *Bluebook's* rule 18 states that online citations are acceptable: (1) "when the information is not available in a printed source; that is, born digital and Internet-only sources" and (2) "if the traditional source is obscure or hard to find and when citation to an electronic source will substantially improve access to the same information contained in the traditional source."²⁴

The American Psychological Association's (APA) style manual is the most widely used style guide in the social sciences. In contrast to the *Chicago Manual*

*of Style*²⁵, and similarly to the *Bluebook*, the APA style guide requires that authors include retrieval dates and other relevant information in order to make future retrieval of the citation easier based on the assumption that this will increase the probability of finding the original source.²⁶ **We have an expectation for this relationship based on our prior research. The logic behind this argument is that authors who have been diligent enough to provide retrieval dates will use extra care when preparing their online references. Providing a retrieval date also suggests that the author is aware that online footnotes tend to lapse. However, this behavioral correlation is not guarantee of permanence.** Thus, our second hypothesis is:

Hypothesis 2: Providing a retrieval date for the citation will be positively related to online citation permanence.

URL Domain

The Internet is structured in a way that allows different browsers on different computer platforms to access and view online information in a similar fashion. The URLs that the browsers display rely on a domain name system (DNS) and the so-called top-level domains or TLDs. According to Wikipedia, the first six generic top-level domains implemented in January 1985 were the following: .com, .edu, .gov, .net, .org, and .mil; the international domain .int was introduced in 1988.²⁷

While past research conclusions about which domain is most stable have varied somewhat, authors generally find that links in the .gov and .org domains seem to be more

stable than .com and .edu domains.²⁸ This was documented, for example, in Wallace Kohler's study.²⁹ Other studies have found similar results. For instance, .gov was the most reliable domain in Tyler and McNeil's study.³⁰ We also found that .gov was the most stable and .com least stable domain.³¹ **The previous two** studies put country-level domains as well as the newer domain names (such as .info, .coop, and so on) in a category for "other." Following this pattern, we formulate the following hypothesis:

Hypothesis 3: Online citations to .gov and .org domains will be more stable than online citations to .com and .edu domains.

URL Structure Level

Tyler and McNeil investigated whether server domain level (i.e., the place of the Web page within the server's directory) influenced online citation failure. They found that Type C domains (defined in their study as links that go to subfiles within a Web site) fared worse than home page links, noting that URLs "with lengthier directory structures and those that specify a particular file name would be the greatest factor that contributes to most web bibliographies rapid obsolescence."³² We also found that first-level links (defined in the study as links to home pages) were more likely to remain accessible over time.³³ Based on these studies, the last hypothesis addresses the relationship between URL level and online citation permanence. We arrive at the following proposition:

Hypothesis 4: Online citation permanence will be negatively related to the URL level of the citation.

There is no consensus in the literature as to whether certain topics or disciplines have more stable online citations. Tyler and McNeil, for example, used the following general topics: arts and humanities, sciences, social sciences, library and information sciences; they did not find any differences between them. Within the area of communication, we wanted to examine whether journals focusing on electronic media have more stable online citations than traditional communication journals.³⁴ In addition to the four hypotheses, therefore, we pose one research question:

Research Question: What is the relationship between journal focus area and online citation permanence in the six communication journals examined in this study?

The next section explains the method used to answer the research question and test the hypotheses.

Method

The study used content analysis methodology to test the above-mentioned hypotheses and answer the research question. The following six communication journals were selected for analysis based on their reputation in the field and high impact factors: *Human Communication Research*, *Journal of Communication*, *Journalism & Mass Communication Quarterly*, *Internet Research*, *Journal of Broadcasting & Electronic Media*, and *New Media & Society*.³⁵ The first three journals were considered traditional in

focus, whereas the other three journals were categorized as focusing on electronic media or new media.

The time frame for the analysis was the publication years 2000, 2001, 2002, and 2003. Following traditional content analysis steps, all research articles published during the four-year period in the six journals were downloaded and saved.³⁶ Editorial notes or book reviews were excluded. All online citations that appeared in the research articles were then entered into a database for analysis. The unit of analysis was the individual URL.

The research articles were downloaded during spring 2004 at a Midwestern land-grant university, using library databases under library copyright agreements through vendors. Trained graduate student coders conducted the coding process. The students were trained on the coding instrument and conducted the initial coding of the content under supervision. Coding was completed during the summer and fall of 2004. The coding categories included several variables pertinent to this study. The first was the permanence of the online citation. If the URL for the citation worked, the variable was coded as 1 (yes); conversely, if the URL did not work, it was coded as 0 (no). Multiple reasons for citation failure exist. Coders recorded the exact error message and then classified it according to type of error: 400 (bad request), 401 (unauthorized access), 403 (forbidden or connection refused by host), 404 (page not found or file not found), 502 (service temporarily overloaded), 503 (service unavailable), or other.³⁷ The year of publication as well as the retrieval date of the online citation, if provided, were also recorded. Furthermore, the URL for each online citation was coded for top-level domain

(TLD) and URL level (for example, home page, second-level link, third-level link, and so on).

Results

The content analysis yielded a total of 1,600 online citations from the six journals published between 2000 and 2003. As expected, their number increased steadily during that period, starting from 276 online citations in 2000, 300 in 2001, 485 in 2002, and 539 in 2003 articles. In other words, the number of online citations in articles from 2000 to 2003 almost doubled. This steady increase is similar to **trends observed in** other disciplines.³⁸

The first objective was to measure the decay rate of online citations in the selected communication journals. When we tested whether the citations were still accessible on the Internet in year 2004, we found that only 1,007 of the 1,600 citations remained accessible while the rest had vanished. In other words, 37 percent of the original URLs were gone, and only 63 percent still worked. This number is comparable to previous studies.³⁹ Furthermore, out of those 1,007 citations that worked, only 588 matched the content of the original source. Only 3.6 percent of the URLs were PDF links (.pdf), whereas 36.9 percent had .html extensions, and 16 percent had .htm extensions. In other words, links to PDF documents were quite rare. The most common error message was the 404 error: page not found.

Our expectation was that journals focusing on new media or electronic media topics would be more likely to rely on online citations than journals with traditional

content focus. Indeed, the three journals focusing on electronic and new media contained a substantially larger number of online citations (1,184) compared with traditional communication journals (416 online citations).

[Table 1. Number of Online Citations per Year]

The four hypotheses were tested using Pearson's chi square comparisons. First, we tested whether the year of publication of the article had a significant impact on online citation permanence. The chi square test showed support for hypothesis 1 (chi square=4.72, $p = .000$, $d.f.=3$). The results indicate a positive relationship, as shown in table 1. The largest number of citations that remained accessible (71.8 percent) was published in articles from 2003, followed by 64.9 percent from articles published in 2002. The percent of accessible citations from 2000 is slightly higher than that for 2001. However, it is important to note that we used publication year as opposed to year of submission or year of acceptance of the article, which would be a better measure of article age. It is desirable for journal editors to make submission and resubmission dates available in the published articles in the future in order to allow more precise estimations.

The second hypothesis stated that providing a retrieval date for the citation would be positively related to online citation permanence. Interestingly, this hypothesis was not supported by the data. More citations that did not provide a retrieval date remained active (66.6 percent) compared with ones that contained retrieval dates (58.1 percent). This indicates that the online medium supercedes citation format and that, contrary to

expectations, adding the date of retrieval of the online citation does not seem to influence its stability.

The third hypothesis predicted that citations referring to .gov and .org domains would be more stable than online citations to .com and .edu domains. This hypothesis was strongly supported (chi square=53.78, $p=.000$, d.f.=4). The cross tabulations show that 71 percent of the citations in the .gov domain and 73 percent of the citations in the .org domain were still active in 2004 (see table 2). They were followed by online citations in the .com domain, which contained 63.9 percent accessible citations. The .edu domain did not fair well since almost half (46.8 percent) of its citations were gone at the time this study was conducted. The lowest number of active citations (49.1 percent) was in the domain labeled "other," which included country-level domains as well as newer domains.

[Table 2. Number of Online Citations per Year by Top Level Domain]

The last hypothesis predicted that URL level would be negatively related to online citation permanence. The chi square test supports this prediction (chi square=71.73, $p=.000$, d.f.=6). The most stable citations (85.2 percent) were located at the top Web level—in other words, they pointed to the home page. The percentage dropped to 63.1 percent for second-level links and 55.9 percent for third-level links that were active. The number of inaccessible links generally increased the further into the Web tree structure.

In summary, three of the four hypotheses were supported. In addition, we also posed one research question about the relationship between journal focus of

communication journals and online citation permanence. To test this relationship, we divided the journals into two groups: traditional and new media journals. While the percentage of working citations from the new media journals group was slightly higher (64 percent) than that of the traditional journals (61 percent), the chi square comparison did not show a statistically significant relationship (chi square=1.08, d.f.=1, n.s.). In other words, the half-life phenomenon has a **similar** impact on both types of journals.

Discussion

Even though some authors argue that "obtaining and printing PDF documents may be more efficient than tracking down the original printed sources,"⁴⁰ this study indicates that half-life remains an issue for top journals in the field of communication. The results of this study show that close to four out of 10 online citations collected from six top refereed journals over a four-year period (2000–2003) do not work. It is important to note that this study examined the most prestigious communication journals, which often have a longer review process, a factor that may have influenced our results. Since we examined only communication journals, it could be profitable to see if our findings hold up in other related fields.

The study also identified some objective characteristics of the URL that can be used by students, researchers, and librarians in order to select more reliable online citations. First, the year of publication of the article can be used a predictor, with more recent dates being more reliable. More recent online sources should, therefore, be preferred.⁴¹ Second, online citations to the .gov and .org domains seem more stable. It

may be desirable to avoid citing links to country-level domains as these are less stable in general. It is also preferable to use top-level links (that is, links to home pages) as these links tend to be more stable. However, this may often contradict the purpose of the citations as home page content is updated often. In general, our results suggest that authors should avoid e-citations to very long URLs.

It is important to note that not all links on the Internet refer to research-related information, of course. Alastair Smith showed that only about 10 to 20 percent of the links even on research sites are analogous to traditional print citations.⁴² Similarly, the majority of the citations in traditional journals still point to print resources. While the relative number of online citations remains small, it is growing steadily. In this study, we did not distinguish between URLs that pointed to journal articles in online databases or URLs to other types of online resources. Another limitation of this study lies in the fact that no attempt was made to locate missing online citations, for example by using online search engines. Previous researchers were able to locate some of the vanished URLs via such methods.⁴³ We also did not account for links that redirected to a new URL (re-routes), which constituted a small, but significant percentage in Tyler and McNeil's study.⁴⁴

Disappearing online information is a problem not only for cited text and data but also for visual content. Larry Gross, John Katz, and Jay Ruby raise the question of data format and archiving issues as well as disappearing images from online newspapers:

Many [images] are **stored** in digital archives, although we may speculate about whether these archives are reliable because, first, they are largely inaccessible to the interested interactor or the critical theorist or historian,

except by purchase, and second, there is no guarantee that picture file formats will not change; older images may one day be as unviewable as data on IBM punch cards or WordStar files on five-and-a-quarter-inch floppies are now. More important, unless one has printed the entire Web page as an integrated unit, it no longer exists at all.⁴⁵

The specter of disappearing footnotes threatens the nature of scholarship and the methods of research across disciplines. Scholarly citations not only disintegrate when URLs vanish; text of original documents can vary because of digital manipulation. As stated earlier, text must be "fixed" in order for footnotes to "prove." There are also vast differences between the printing press and the Internet, especially when it comes to printed documents like books. Bugeja calls the book an "ultimate fire-walled medium," with exact printed copies distributed to libraries.⁴⁶ When one rips or marks up pages on library books, that person commits a crime because other scholars might need the work for future reference. That is not the case with the Internet, which allows a user to select, copy, paste, and manipulate original works and then post them online or file them in a database.

The standards of scholarship continue to require fixed text that scholars can access in original form and formats, otherwise the promise of digital data as embraced by such entities as the National Science Foundation, which hopes to fix format so that information is assessable technologically in the long term, will fall short of its goal simply because critical online citations will have lapsed.⁴⁷ In *Defenders of the Text: The Traditions of Scholarship in an Age of Science*, Anthony Grafton notes, "Only systematic

comparison between a given work under analysis and many earlier and contemporary texts can make the modern reader familiar with the inherited technical language a past scholar used; without that familiarity one cannot distinguish between the novel and the traditional, the original and the traditional.⁴⁸ Hence, the task to preserve scholarship as evolved since Leopold von Ranke entails the stabilization of a dynamic medium, the Internet, whose alluring features are easy access to data and manipulation thereof—**the opposite of printed text in physical libraries.**

In Search of Solutions

The half-life phenomenon has no easy solutions. As far as individual citations are concerned, stabilization is difficult in part because it remains unclear what factors affect online citation permanence. Some have focused on implementing technical solutions to the moving URL problem by introducing the so-called digital object identifier (DOI) system. Digital object identifiers should alleviate the management of information in a digital environment.⁴⁹ Others have suggested user-focused, short-term solutions such as saving cited online documents on the desktop and printing hard copies of the cited URLs.⁵⁰

Individual institutions are aware of the half-life issues and have taken steps to counter link rot. At the University of Rochester River Campus, an electronic archive creating "DSpace" (based on a system developed by the Massachusetts Institute of Technology) "allows academic departments to build e-real estate where faculty can store their documents."⁵¹ Susan Gibbson, director of digital library initiatives, states that a key

feature of DSpace is "persistent URLs"—an attraction that faculty members dealing with the half-life of Internet citations may find particularly attractive.⁵² The library at the University of Pennsylvania uses a system called PennText to digitize online journals and promotes DOIs as "a bar code for intellectual property," ensuring that a registered URL will permanently exist.⁵⁶⁵³

At present, however, no universal archiving system exists to stabilize online text whose access is becoming increasingly popular—even though researchers increasingly encounter the erosion of online citations. Indeed, Davis argues that convenience may play a role in the increasing popularity of online citations.⁵⁴ In addition, scholars studying online phenomena may need to quote an Internet source merely because such sources are germane to the subject matter of their research. In sum, **the very researchers who have a vested interest in stabilizing the Internet or finding new research methods to adjust for its dynamic nature** are, ironically, the ones whose scholarship may be most affected by the half-life of Internet footnotes. It seems likely that future scholars who need to re-acquire lost online sources will have to look in archives such as the Internet Archive or the Wayback Machine to locate missing sources. Such archives, whose goal is to create backups of the World Wide Web, **may** be more reliable than online search engines such as Google. Future research should compare the effectiveness of these two methods for retrieving vanished Internet citations.

Conclusions

To continue the influence of Bacon and Locke and the scientific tradition, three objectives must be met:

- *Text must be stabilized.* Researchers must have access to original documents rather than manipulated versions thereof.
- *Citations in those documents must be retrievable.* Researchers must have access to archives that guarantee both the longevity of technical formats and cited URLs.
- *The digital library must remain a repository of fact in addition to being a dissemination point of information.* The emphasis on the latter has eroded the former so that no universal archive exists to ensure the permanency associated with the printing press and the scientific tradition.

Thus, the time-honored recommendation, "consider the source," requires reconsideration. The maxim has been a part of rhetorical history, referring to the reliability of the person or data being cited in a work. That standard still remains. However, in the digital era, scholars also must consider the online source that generates such a reference. The year of publication, the **Web site domain, and the URL** level are important indicators of the likelihood of a site's availability. Otherwise, the historic standard is undermined by the technological one. Toward that end, despite the lack of convenience, access to printed material will remain the most reliable source for citation.

Daniela V. Dimitrova is vice head and program chair, Communication Technology Division, **Association for Education in Journalism and Mass Communication (AEJMC)**, and assistant professor, Greenlee School of Journalism and Communication, Iowa State University, Ames, IA; she may be contacted via e-mail at danielad@iastate.edu.

Michael Bugeja is director, Greenlee School of Journalism and Communication, Iowa State University, Ames, IA; he may be contacted via e-mail at bugeja@iastate.edu.

Notes

¹ Mary K. Taylor and Diane Hudson, "'Linkrot' and the Usefulness of Web site Bibliographies," *Reference & User Services Quarterly* 39, 3 (2000): 273–80; David C. Tyler and Beth McNeil, "Librarians and Link Rot: A Comparative Analysis with Some Methodological Considerations," *portal: Libraries and the Academy* 3, 4 (2003): 615–32.

² Kenneth Burke, *A Rhetoric of Motives* (Berkeley, CA: University of California Press, 1969).

³ Anthony Grafton, *The Footnote: A Curious History* (Cambridge, MA: Harvard, 1997).

⁴ Patricia Bizzell and Bruce Herzberg, *The Rhetorical Tradition: Readings from Classical Times to the Present* (New York: St. Martins, 1990), 504.

⁵ *Ibid.*, 10.

⁶ Ibid.

⁷ Ibid., 697.

⁸ According to Bizzell and Herzberg, the philosophers of the Enlightenment "reconsidered the source and status of knowledge, paying particular attention to the psychological processes of perception, reflection, and communication in an attempt to determine how it was possible to discover the truths within the physical world that were so important for science's progress" (p. 504).

⁹ Grafton, 34.

¹⁰ Marcus A. Banks, "The Excitement of *Google Scholar*, the Worry of *Google Print*," *Biomedical Digital Libraries* 2, 1 (2005): 1–3.

¹¹ Michael Bugeja and Daniela V. Dimitrova, "Exploring the Half-Life of Internet Footnotes," *Iowa Journal of Communication* 37, 1 (2005): 77–86; Philip M. Davis, "Effects of the Web on Undergraduate Citation Behavior: Guiding Student Scholarship in a Networked Age," *portal: Libraries and the Academy* 3, 1 (2003): 41–51; Mary Rumsey, "Runaway Train: Problems of Permanence, Accessibility, and Stability in the Use of Web Sources in Law Review Citations," *Law Library Journal* 94, 1 (2002): 27–39; John Markwell and David W. Brooks, "'Link Rot' Limits the Usefulness of Web-based Educational Materials in Biochemistry and Molecular Biology," *Biochemistry and Molecular Biology Education* 31, 1 (2003): 69–72; Carmine Sellitto, "The Impact of Impermanent Web-located Citations: A Study of 123 Scholarly Conference Publications," *Journal of the American Society for Information Science and Technology* 56, 7 (2005): 695–703; Taylor and Hudson; and Tyler and McNeil.

¹² Markwell and Brooks.

¹³ Jeffrey D. Kushkowski, "Web Citation by Graduate Students: A Comparison of Print and Electronic Theses," *portal: Libraries and the Academy* 5, 2 (2005): 259–76.

¹⁴ Philip M. Davis and Suzanne A. Cohen, "The Effect of the Web on Undergraduate Citation Behavior 1996–1999," *Journal of the American Society for Information Science and Technology* 52, 4 (2001): 309–14.

¹⁵ Ibid.

¹⁶ Bugeja and Dimitrova; Wallace Koehler, "A Longitudinal Study of Web Pages Continued: A Report after Six Years," *Information Research* 9, 2 (2004), <http://InformationR.net/ir/9-2/paper174.html> (accessed March 31, 2006); Tyler and McNeil.

¹⁷ Davis and Cohen; Kushkowski.

¹⁸ Paul O. Jenkins, "They're Not Just Using Web Sites: A Citation Study of 116 Student Papers," *College and Research Libraries News* 63, 3 (2002): 164.

¹⁹ Daniela V. Dimitrova and Michael Bugeja, "The Half-Life of Internet References Cited in Communication Journals," *New Media & Society* (forthcoming).

²⁰ Rumsey.

²¹ Tyler and McNeil.

²² Dimitrova and Bugeja, "The Half-Life of Internet References."

²³ Rumsey.

²⁴ Sheryl Nyberg, "Cite-Checking and Library Research," M.G. Gallagher Law Library, University of Washington School of Law, <http://lib.law.washington.edu/ref/citecheck.html> (accessed March 31, 2006).

²⁵ **However, the Chicago Manual of Style 17.12 states that for fields that are very "time sensitive" such as law and medicine, the access date is advisable to include.**

²⁶ Citation StyleGuide Handbook, Manuscripts: Guidelines for Preparing your Paper, <http://citationonline.net/CitationHelp/csg04-manuscripts-apa.htm> (accessed March 31, 2006).

²⁷ Wikipedia: The Free Encyclopedia, s.v. generic top level domain, http://en.wikipedia.org/wiki/Generic_top-level_domain (accessed March 31, 2006).

²⁸ See Tyler and McNeil's discussion for details.

²⁹ Koehler.

³⁰ Tyler and McNeil.

³¹ Dimitrova and Bugeja, "The Half-Life of Internet References."

³² Tyler and McNeil, 626.

³³ Dimitrova and Bugeja, "The Half-Life of Internet References."

³⁴ Tyler and McNeil.

³⁵ According to Thompson Scientific, the company that publishes *Journal Citation Reports (JCR)*, impact factors work as follows: "The *JCR* provides quantitative tools for ranking, evaluating, categorizing, and comparing journals. The impact factor is one of these; it is a measure of the frequency with which the 'average article' in a journal has

been cited in a particular year or period. The annual *JCR* impact factor is a ratio between citations and recent citable items published. Thus, the impact factor of a journal is calculated by dividing the number of current year citations to the source items published in that journal during the previous two years,"

<http://scientific.thomson.com/free/essays/journalcitationreports/impactfactor/> (accessed March 31, 2006).

³⁶ Daniel Riffe, Steven Lacy, and Fred. G. Fico, *Analyzing Media Messages: Using Quantitative Content Analysis in Research* (Mahwah, NJ: Lawrence Erlbaum Associates, 1998).

³⁷ See Brown University's site,

http://www.brown.edu/Departments/Taubman_Center/errormsg.html, for more details about error messages (accessed March 31, 2006).

³⁸ Rumsey.

³⁹ Bugeja and Dimitrova, "Exploring the Half-Life of Internet Footnotes;" Tyler and McNeil.

⁴⁰ Nyberg.

⁴¹ Even though we measured only the year of publication of the article (and not the actual URL), we can assume that more recent articles cite more recent Web pages.

⁴² Alastair Smith, "Web Links as Analogues of Citations," *Information Research* 9, 4 (2004), <http://informationr.net/ir/9-4/paper188.html> (accessed March 31, 2005).

⁴³ Taylor and Hudson.

⁴⁴ Tyler and McNeil.

⁴⁵ Larry Gross, John S. Katz, and Jay Ruby, *Image Ethics in the Digital Age* (Minneapolis: University of Minnesota Press, 2003), 19.

⁴⁶ Michael Bugeja, "Such Stuff as Footnotes are Made On," *Inside Higher Ed* (April 22, 2005), <http://www.insidehighered.com/views/2005/04/22/bugeja> (accessed March 31, 2005).

⁴⁷ Doug Lederman, "Delivering on the Promise of Digital Data," *Inside Higher Ed* (October 13, 2005), <http://www.insidehighered.com/news/2005/10/13/digital> (accessed March 31, 2005).

⁴⁸ Grafton, 13.

⁴⁹ See <http://www.doi.org> for details (accessed March 31, 2006).

⁵⁰ Bugeja and Dimitrova.

⁵¹ Bill Meehan, "Going Digital," *SLIS Network* (2005): 10.

⁵² *Ibid.*, 11.

⁵³ *Ibid.*

⁵⁴ Davis.