June 2008

Measuring the "Google Effect" at JSTOR

Bruce Heterick

JSTOR, bruce.heterick@jstor.org

Follow this and additional works at: <https://docs.lib.purdue.edu/atg>

Part of the [Library and Information Science Commons](#)

Recommended Citation

Heterick, Bruce (2008) "Measuring the "Google Effect" at JSTOR," *Against the Grain*: Vol. 20: Iss. 3, Article 2.

DOI: <https://doi.org/10.7771/2380-176X.2394>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

Measuring the “Google Effect” at JSTOR

by **Bruce Heterick** (Director, Library Relations, JSTOR, ARTstor, Portico, Aluka, 149 Fifth Avenue, New York, NY 10010; Phone: 212-358-6416; Fax: 212-358-6499) <bruce.heterick@jstor.org> www.jstor.org www.artstor.org
www.portico.org www.aluka.org

Someone once told me that statistics are like bathing suits: what they reveal is interesting, but what they conceal is essential. Over the past 24 months, since **Google** began indexing the journal content preserved in the **JSTOR** archive and making that content discoverable in **Google** and **Google Scholar**, we have had the opportunity to gather usage, access, and linking statistics. These statistics reveal the changing dynamics of content discovery, and provide insights into how faculty, students, and scholars will use **JSTOR** for their future research and pedagogical needs. We have also witnessed the “**Google Effect**.” **JSTOR** has had organizational, operational, and strategic effects that statistics do not always reveal.

By almost any statistical measure, use of the **JSTOR** archive has grown at a rapid pace since the introduction of the first **JSTOR** collection in 1997.

JSTOR Significant Accesses | 2001-2007

FIGURE I

In 2007, there were over 500 million significant accesses to the **JSTOR** archive (an increase of 38% from 2006), including over 137 million articles viewed. (See **FIGURE I**.) There are many contributing factors to this growth in usage: an increasing number of participating institutions (~4,500); an increasing number of titles available in the archive (~800); as well as a general increase in the availability (and acceptance) of digital content in the past decade. There has also been an appreciable increase in the link referrals that **JSTOR** has received over the years from resources with whom we have a designated linking relationship (e.g., Research Papers in Economics (RePEc), History Coop, MathSciNet, Cross-Ref, SFX, etc.). **JSTOR** has approximately 38 such relationships at this point, and in 2007, they drove over 6.5M links to **JSTOR** (See **FIGURE II**).

Rumors from page 32

Got a **Mother's Day** email from **Farfuri Xhaja** of Albania! Remember **Farfuri**? She was a **Boggle scholar** who came to Charleston from Albania. We interviewed her in **ATG** in issue v.8 no.5? **Farfuri** is doing well. She is still working in the **Albanian National**

continued on page 54

**JSTOR Inbound Links from Linked Partners
2001-2007**

FIGURE II

Successful links into **JSTOR** from these partners increased by 27% from 2006 to 2007, following a 23% increase from 2005 to 2006. Those sizeable increases, however, are dwarfed by the increase in links from “unknown” sources (links from requestors who do not have a formal linking agreement with **JSTOR**, and therefore, we can't discern their exact origin). Typically, these are links from library OPACs, faculty course syllabi, and other Websites that have captured the stable URL link for a **JSTOR** article, but have not been assigned an “origin parameter” by **JSTOR**. From 2005 to 2006, the number of links from this category grew by 581% to over 23 million; and in 2007, that number grew again by 117% to over 50 million (See **FIGURE III**). Research into the **JSTOR** Weblogs attributes this marked increase to: (a) libraries that have incorporated a **Google** search box in their library homepage or portal; (b) Internet Service Providers (e.g., **AOL**, **Comcast**) that have done the same; and (c) **Google** crawling Websites that have a substantial number of **JSTOR** links (e.g., **RePEc**).

**JSTOR Inbound Links from Partners and “Unknown”
2001-2007**

FIGURE III

As you can see, this “**Google Effect**” changes the scale of links from the “Unknown” origins by an order of magnitude, as compared to the links from **JSTOR**'s designated linking partners. Incredibly, another order of magnitude change in scale is introduced when we begin to look at the

continued on page 46

Measuring the “Google Effect” from page 44

number of links coming to JSTOR directly from Google and Google Scholar (See FIGURE IV). In early 2006, in response to students, faculty, and researchers using the JSTOR archive, JSTOR finalized an agreement with Google to allow the popular search engine to begin “crawling” and indexing book reviews and full-length articles archived in JSTOR for discovery purposes.¹ Those at participating institutions may access the articles seamlessly in JSTOR by clicking on the URL provided in the Google/Google Scholar search result. (Off-site users must first authenticate via their libraries’ remote access system.) This route to JSTOR has become increasingly popular: in 2007, almost 175 million referrals from Google/Google Scholar resulted in just over 32 million links to articles in JSTOR (23% of total article views in the archive). The number of links from Google-referring URLs increased by 159% from 2006 to 2007.

Total JSTOR Inbound Links, including Google
2001-2007

FIGURE IV

So, what have been the positive aspects of this “Google Effect” for JSTOR? The most obvious positive impact is that students, faculty, and researchers at participating institutions can now get to the content in JSTOR from the place they have chosen to start their research: the Google and Google Scholar search box. While www.jstor.org continues to grow as a destination for research — particularly in certain disciplines — the number of links coming to JSTOR from Google-referring URLs is an increasingly larger component of the significant accesses in JSTOR. Another positive impact has been the increased exposure of the journals archived in JSTOR to a broader audience than could ever be reached by JSTOR itself. We hypothesize that this exposure, in turn, has begun to have an impact on the most used disciplines and journals in JSTOR. If one were to compare usage by discipline in JSTOR (2004 vs. 2007), while also looking at the top journals being referred from Google/Google Scholar, some interesting trends begin to emerge (See FIGURE V).

Top Disciplines Accessed in JSTOR | 2004 vs. 2007

2004		2007	
Disciplines	Accesses	Disciplines	Accesses
1. History	16,566,200	1. Lang & Lit.	31,833,045
2. Economics	14,552,178	2. Bus/Econ	29,181,793
3. Political Sci	12,599,486	3. History	24,510,272
4. Lang & Lit.	10,738,228	4. Political Sci.	18,671,356
5. Sociology	10,305,708	5. Sociology	16,021,218
6. Ecology	6,227,587	6. Education	12,640,714
7. Business	5,456,328	7. Biological Sci	12,303,144
8. Anthropology	4,987,091	8. Art & Art Hist.	9,412,240
9. Gen. Science	4,761,610	9. Anthropology	9,386,176
10. Philosophy	4,363,140	10. Law	9,352,744

FIGURE V.

A closer look at the 2007 discipline-based usage at JSTOR shows the emergence of Education (#6), Biological Sciences (#7), Art/Art History (#8), and Law (#10). JSTOR has added a number of titles in those disciplines since 2004 — in Education and Biological Sciences in particular — so it is not particularly surprising to see those disciplines move into

the top 10. It is a little more surprising to see the Art/Art History and Law disciplines enter the top disciplines accessed in JSTOR in 2007. Could the usage growth in these disciplines be a by-product of the Google indexing? Perhaps. In looking at the usage logs to identify which journals in JSTOR are most commonly linked to from Google-referring URLs, there are a number of art/art history and law titles in the top 100. Further analysis is required to determine the exact correlation, but it is no surprise that “discoverability” via Google/Google Scholar has brought the journal content in those disciplines to a wider audience than they might normally have reached.

What are the primary challenges resulting from the “Google Effect” for JSTOR? The foremost challenge is *getting people to the information that they want*. For the 183 million referrals² that could not be authorized or authenticated — originating from independent researchers, non-participating institutions, or from participating institutions wishing to access back issues in a JSTOR collection to which they were not licensed — access was essentially denied.³

This is problematic for a not-for-profit organization that has, as part of its mission, the desire to extend access to the archive as broadly as possible. This increased exposure to these unaffiliated/unauthenticated users has put significant pressure on JSTOR to deploy access options to meet this demand. The aforementioned Publisher Sales Service is one thread of a response, and offering access to the JSTOR archive in the for-profit community is another.⁴ But this issue requires a multi-threaded response that offers scalable solutions. How does JSTOR improve this user experience and extend its mission without devolving the economic model that sustains the organization?

Another challenge JSTOR has to be concerned with, is additional cost of *enhanced discovery*. It would probably come as no surprise to many of you who have been dealing with the “Google Effect” at your own institutions that the costs associated with this scale of increased usage are very real. In 2006, for instance, user support inquiries from unauthenticated/unaffiliated users increased by 500%. Total user support inquiries increased by 110% in 2006 and 86% in 2007. In addition, the increased infrastructure costs to handle the massive increase in Google-driven traffic have been significant. As a resource that is expected to be available 24/7/365 with reasonable response times has been a challenge. Finally, it is important to remember that the impact we have been discussing has been the result of one (albeit HUGE) discovery relationship. As JSTOR is asked to consider other discovery platforms (e.g., Microsoft Live Search, OCLC WorldCat, search engines in other countries), the cost issues are only amplified further.

In conclusion (the finest phrase in the English language), we are just beginning to see institutions, publishers, and providers begin to measure the “Google Effect” and distill usable lessons from those experiences. To paraphrase Catherine Aird, for some, JSTOR may provide a useful example, and for others, JSTOR may offer a horrible warning.⁵ Regardless of which camp you happen to fall into, benign neglect to the reality of the changes that Google and its brethren are having on the way in which students, faculty, and researchers interact with online content & the tools that make it available would be a mistake.

As authors Stan Davis and Jim Botkin opined in their book, *The Monster Under the Bed*:⁶ “Every time the infrastructure shifts, everything else shifts with it.” The infrastructure has shifted, folks. Nostalgia isn’t what it used to be. 🐉

Endnotes

1. “JSTOR’s ‘Google Effect,’” *JSTORNEWS*, March 2007, No. 11, Issue 1; http://news.jstor.org/jstornews/2007/03/march_2007_no_11_issue_1_jstor.html.
2. 142 million from Google-referring URLs and 41 million from “Unknown” sources, but attributable to Google/Google Scholar.
3. From the User Response Page, JSTOR typically offers the user access to the first page of the article (for context purposes), as well as options to: (a) identify JSTOR participating libraries (in case they might have access privileges), (b) view detailed publication information (should he/she desire to contact the publisher directly), and (c) purchase the article from the publisher (if the publisher has opted to offer that service). In 2007, over 35,000 articles were purchased.
4. Since its inception, JSTOR has only offered access to not-for-profit institutions. In June 2008, JSTOR will begin offering access to for-profit institutions through its Corporate & For-Profit Access Initiative.
5. The quote attributed to Catherine Aird is: “If you can’t be a good example, then you’ll just have to be a horrible warning.”
6. Davis, Stan and Jim Botkin. *The Monster Under the Bed*. Touchstone, New York, NY 1994.