
 **WHITE PAPER**

The Sources in Student Writing – Secondary Education

Sources of Matched Content and Plagiarism in Student Writing



Prevent Plagiarism.
Engage Students.

www.turnitin.com

Table of Contents

1.0 Summary 3

2.0 About This Study..... 4

3.0 Categorizing Student Sources 5

4.0 The Most Popular Sources and Sites 6

5.0 The Educational Value of Student Sources..... 7

6.0 Paper Mills & Cheat Sites Continue to Thrive 8

7.0 Comparing Secondary and Higher Ed Sources 9

8.0 Recommendations for Educators..... 10

About Turnitin 11

1.0 Summary

This study examines which Internet sources secondary education students use in written work and the implications of their choices. It is based on an analysis of over 44 million content matches from more than 9 million student papers submitted to Turnitin between July 2011 and June 2012.

KEY FINDINGS INCLUDE:

Students are relying on sources that have weak academic validity. 50% of matches lead to sites that are academically suspect, including cheat sites and paper mills, shopping sites, and social and user-generated content.

Most troubling, 18% of content matches come from paper mills and cheat sites. In looking at the issue of plagiarism, it is safe to assume, at minimum, 18% of matched content comes from sources of academic disrepute.

Secondary student sources vary considerably from the sources higher education students use in their writing. Not surprisingly, the quality of sources in secondary student writing is lower than those of higher education students, primarily because of a greater reliance on Social Networking & Content Sharing sites and Shopping sites.

Educators should consider taking time to teach students how search engines work and how to evaluate and judge the quality of online content. The wealth of information on the Internet is making the process of searching, discovering, processing and communicating information a lost art. Instructors should consider assignments that will help students understand how search engines work as well as how to review the information that they find online.

2.0 About This Study

This paper is an update to a report published in November 2011, entitled “Plagiarism and the Web: A Comparison of Web Sources for Higher and Secondary Education Students.”¹

Papers submitted to the Turnitin originality checking and online grading service are compared against 30 billion current and archived web pages, 120 million textbook and academic journals and 275 million student papers. For the purposes of this study, only matched content between student papers and Turnitin’s archive of the Web were considered.

For each month of the study, Turnitin documented the top 100 Internet sites (domains) that displayed matches between student papers and Turnitin’s database of 30 billion web pages. Each site was identified, reviewed and placed into one of six categories:

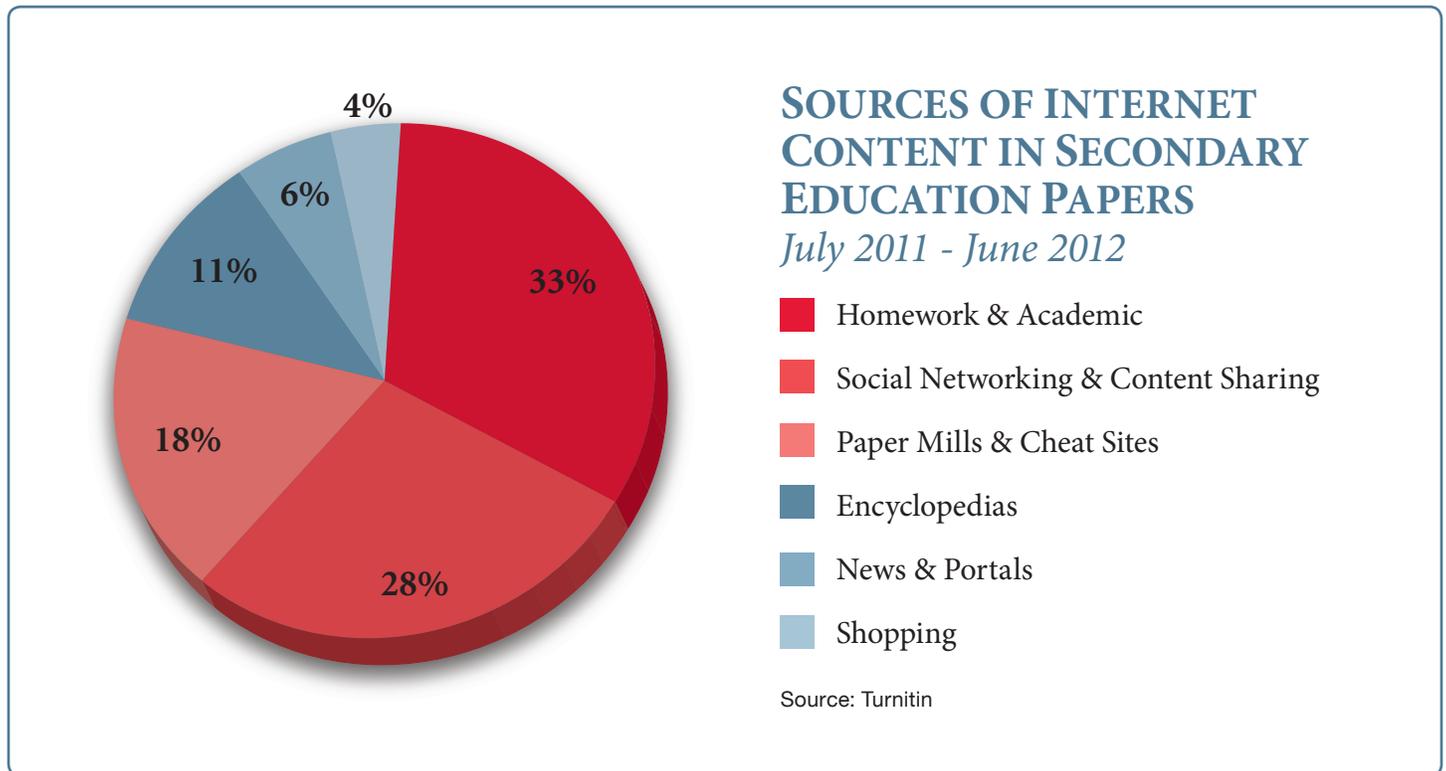
Category	Definition	Examples
Social Networking & Content Sharing	Sites that rely on user-generated content rather than professionally-published content.	Answers.com, Facebook, Scribd, SlideShare, Yahoo Answers
Homework & Academic	Academic, educational and homework help sites that offer a range of content to educate students.	Google Books, Project Gutenberg, Mayo Clinic, MedLibrary, Springer, Sage
News & Portals	This category consists of professionally-published content.	Associated Content, Examiner, The Huffington Post, The New York Times
Paper Mills & Cheat Sites	Sites that promote the sale or sharing of written work.	All Free Papers, All Free Essays, Essay Mania, Frat Files, Other People’s Papers
Encyclopedias	Reference and summary sites	Wikipedia, Britannica, Encyclopedia.com, Internet Movie Database
Shopping	Sites that offer reviews and sell products	Amazon, Barnes & Noble, Buy.com

Note: The matched texts identified for study in this report **may or may not have** been properly cited and attributed. Also, no school, student or other identifying information was viewed when conducting the analysis.

¹See: http://pages.turnitin.com/plagiarism_sec_vs_he.html

3.0 Categorizing Student Sources

The chart below shows the percentage of sources by category of student writing based on six categories.



Category	2011-2012	2010-2011	Change
Homework & Academic	33%	23%	+10
Social Networking & Content Sharing	28%	31%	-3
Paper Mills & Cheat Sites	18%	14%	+4
Encyclopedias	11%	11%	0
News & Portals	6%	12%	-6
Shopping	4%	5%	-1

The most notable changes in the results paint a mixed picture of student source use. On the positive side, there was a ten-percentage point increase in student use of Homework and Academic sites. Also, student reliance on Social Networking and Content Sharing web sites showed a three-percentage point decrease compared to July 2010 through June 2011. Unfortunately, this latter improvement was offset by an increase in the use of Paper Mills and Cheat Sites, which increased four percentage points. Secondary students relied on Shopping sites such as Amazon.com and Barnes & Noble 4% of the time, most likely for information from reviews about books and movies.

4.0 The Most Popular Sources and Sites

MOST POPULAR SITES.

The top 10 most popular Internet sources for student writing include four Social Networking & Content Sharing sites, three Paper Mills & Cheat Sites, two Homework & Academic sites and one Encyclopedias site.

2011-2012 Top Sites	%	Category	2010-2011 Top Sites	%	Category
wikipedia.org	8%	Encyclopedias	wikipedia.org	8%	Encyclopedias
answers.yahoo.com	7%	Social	answers.yahoo.com	8%	Social
enotes.com	3%	Homework	answers.com	3%	Social
answers.com	3%	Social	enotes.com	3%	Homework
oppapers.com	3%	Paper Mills	slideshare.net	2%	Social
scribd.com	3%	Social	scribd.com	2%	Social
slideshare.net	2%	Social	oppapers.com	2%	Paper Mills
essaymania.com	2%	Paper Mills	essaymania.com	2%	Paper Mills
shmoop.com	2%	Homework	amazon.com	2%	Shopping
medlibrary.org	2%	Homework	123helpme.com	2%	Paper Mills

Wikipedia remains the top source of unoriginal writing in student papers. The popular, crowd-sourced Encyclopedia site comprised 8% of all matches, which is the same percentage as the previous year.

Eight of the top 10 sites appear in our most recent study as well as in the previous study (July 2010 to June 2011). Two Paper Mills & Cheat Sites appear in the top 10 list of the most recent academic year compared to three in the previous year.

In terms of potential for educational value, four sites—Wikipedia, Enotes.com, Shmoop and MedLibrary—could be considered to have academic merit. Two sites share (and/or) sell written work (Oppapers and EssayMania), thus offering little to no educational value. The other four sites all fall under the Social Networking and Content Sharing category – SlideShare, Scribd, Yahoo! Answers and Answers.com. The educational value of the content in these sites is quite varied, but instructors would most likely prefer not to see these types of sites included in a list of cited works.

5.0 The Educational Value of Student Sources

While results may vary by individual sites, Homework & Academic sites, Encyclopedias and News & Portal Sites all provide some degree of content with educational value. Given these criteria, 50% of sources in student writing come from sites that contain educational value.

Category	2012
Homework & Academic	33%
Encyclopedias	11%
News and Portals	6%
Total	50%

Conversely, Paper Mills & Cheat Sites, Shopping sites and Social Networking & Content Sharing sites fall significantly below the bar of offering educational value. Book reviews (from Amazon.com, for example, the most popular domain in Shopping) and anonymous opinions on answer sites (Answers.com, for example) also do not pass muster as material that should be included in student writing. Given these criteria, 50% of sources that appear in student writing may be considered inappropriate for academic work.

Category	2012
Social Networking & Content Sharing	28%
Paper Mills & Cheat Sites	18%
Shopping	4%
Total	50%

6.0 Paper Mills & Cheat Sites Continue to Thrive

The category of Paper Mills & Cheat Sites provides concrete examples of sources students should not use in their writing. Unfortunately, these sites continue to be a common source for materials that appear in student papers, growing from 14-18% in the most recent academic year.

While educators can debate the educational value of using Wikipedia or SlideShare.net in research, it is difficult to imagine a scenario where content from sites such as Opppapers.com (i.e. Other People's Papers) or Fratfiles.com would be academically permissible.

Below is a list of the top 10 cheat sites and paper mills that appear in student work.

Top Cheat Sites SE 2012		# of Matches
1	oppapers.com	1,245,097
2	essaymania.com	932,709
3	antiessays.com	798,557
4	123helpme	794,807
5	allfreeessays.com	604,413
6	bignerds.com	534,955
7	slashdoc.com	526,152
8	customessaymeister.com	525,242
9	novelguide.com	410,964
10	termpaperslab.com	302,811

These sites, along with a number of other cheat sites and paper mills that appear further down in the list represent a troubling 18% of all sources in student writing. To put it bluntly, at least 18% of all sources used in secondary education papers are plagiarized.

7.0 Comparing Secondary and Higher Ed Sources

A comparison of higher and secondary sources shows some similarities but also stark differences.

Secondary Sources	%	Higher Ed Sources	%
Homework & Academic	33%	Homework & Academic	33%
Social Networking & Content Sharing	28%	Social Networking & Content Sharing	23%
Paper Mills & Cheat Sites	18%	Paper Mills & Cheat Sites	19%
Encyclopedias	11%	Encyclopedias	14%
News & Portals	6%	News & Portals	10%
Shopping	4%	Shopping	1%
Total 100%	100%	Total	100%

50% of the sources secondary students use in writing comes from legitimate academic resources (Homework & Academic, Encyclopedias and News & Portals), while higher education students relied on these sites 57% of the time. A closer look at the categories shows where secondary and higher education students differ.

- Both levels show the same reliance on Homework & Academic sites (33%), the most popular category.
- Secondary students depend more on Social Networking & Content Sharing sites, most likely because they have less experience assessing the value of student sources.
- Reliance on Paper Mills & Cheat Sites remains stubbornly high though roughly equal for both categories.
- Higher education students showed a great dependence on Encyclopedias sites (largely Wikipedia) and for News & Portal sites that present professionally-published content.
- Secondary students favor Shopping sites to a much great degree than higher education students, where sites such as Amazon.com and Barnes & Noble barely register. Once again, this is most likely a reflection of the difficulty younger students face in assessing and including academically-viable source material.

8.0 Recommendations for Educators

There are four steps educators can take to ensure students use and cite the right sources on the Web.

1. DON'T ASSUME STUDENTS KNOW HOW TO CONDUCT PROPER RESEARCH

Based on this analysis, it is clear that many secondary education students do not have a strong grasp of what it means to conduct proper academic research. Instructors who assume that students know how to sift through the myriad sources on the Internet and find educationally valuable information will most likely find papers with spurious sources on their desks at the end of the term. In writing intensive courses, instructors should dedicate their first class and assignments to educating students on what constitutes academically-appropriate sites and share best practices for uncovering sources that have high educational value.

2. GOOGLE IS THE FIRST STEP, NOT THE LAST, IN RESEARCH

Students think they know how to use Google, but when it comes to academic research, it is clear that they do not. Google's algorithms are tuned for relevance and popularity, not academic authority. When it comes to finding academic sources, teach students that they cannot defer to Google results and instead must use their judgment to decide what sources are of academic value. For more information, please view the webcast "*New Demands, New Approaches: Achieving Information Literacy and Competency in the 21st Century.*" http://pages.turnitin.com/HE_22_archive.html

3. DESIGN ASSIGNMENTS THAT MAKE STUDENTS RESEARCH THE PRE-INTERNET WAY

As the Pew Research Center described recently in a study, research in the digital age has evolved from "a relatively slow process of intellectual curiosity and discovery to a fast-paced, short-term exercise aimed at locating just enough information to complete an assignment."² Instructors should consider creating assignments where no citations can come from the Internet. While artificial, these types of assignments would help students appreciate the more deliberative pace that is required when searching for proper sources in academic writing.

4. EMPLOY TOOLS LIKE TURNITIN

In this digital age, educators need to use educational technology to better assess the value of student sources while providing better feedback to students. Turnitin helps instructors to uncover improper sources in student papers and to give better feedback on student writing. Within the first year of usage, Turnitin users see a reduction in unoriginal content of 10-35%. By the third year, institutions see levels of unoriginality fall by 50%, with some seeing improvements up to 75%. Moreover, educators who use Turnitin report a 46% increase in student engagement, a 52% increase in the quality of feedback to students and a 31% reduction in time spent grading.

²Pew Research Center, CollegeBoard, and National Writing Project, *How Teens Do Research in the Digital World* (Washington DC: PIP, 2012) 3.

About Turnitin

Turnitin is the global leader in evaluating and improving student writing. The company's cloud-based service for originality checking, online grading and peer review saves instructors time and provides rich feedback to students. One of the most widely distributed educational applications in the world, Turnitin is used by more than 10,000 institutions in 126 countries to manage the submission, tracking and evaluation of student papers online. Turnitin also offers iThenticate, a plagiarism detection service for commercial markets, and WriteCheck, a suite of formative tools for writers.

Turnitin is backed by Warburg Pincus and is headquartered in Oakland, Calif., with an international office in Newcastle, U.K.

For more information, please visit <http://turnitin.com>.

Turnitin is a service of iParadigms, LLC
1111 Broadway, 3rd Floor
Oakland CA, 94607 USA

USA/Canada: 866-816-5046
International: +1 510-764-7600
sales@turnitin.com
www.turnitin.com

All products and services mentioned in this document are trademarks or registered trademarks of their respective owners.
©2013 iParadigms, LLC All rights reserved. Version 0113.