THE HENRY J. KAISER FAMILY FOUNDATION

See No Evil: How Internet Filters
Affect the Search for Online
Health Information

sexual health xxx safer sex

sexual health

health chlamydia

sex education

sex guide

INTERNET

AIDS

condoms

breast cancer

topless

safer sex

sex guide condoms

chlamydia sex education

sexual AIDS

health STDs

xxx

safer sex FILTERING

pregnancy

sex quiz

sex education

birth control

condoms

porn

By Victoria Rideout, MA, Caroline Richardson, Ph.D., and Paul Resnick, Ph.D.

Contents

Introduction

Concern has been raised as to whether Internet filters block young people's access to non-pornographic health information. Despite prolonged and impassioned debate about the potential impact of blocking software, there is surprisingly little empirical evidence regarding blocking errors, particularly as they might affect access to health information. Most reports have been anecdotal rather that systematic in nature and have not focused on health searches in particular.

The study reported here is a large-scale, scientific study designed to help determine whether Internet filters are likely to block young people's access to non-pornographic health information. The study simulates young people's online health information searches and measures the impact of seven different filtering products on those searches, looking at both the effectiveness of the filters at blocking pornography and the rate at which they also block non-pornographic health information.

The basic design of the study is to:

> Run searches on a variety of youth-oriented health topics on six of the search engines that are most popular among young people;

> Test the thousands of health URLs resulting from these searches against six different filtering products commonly used in schools and libraries, and one product widely used in the home; and

> Run searches on a variety of porn search terms, and test the hundreds of URLs resulting from those searches against the same filtering products.

Importantly, most filtering software allows system administrators to specify blocking configurations, giving individual schools and libraries the ability to tailor the blocking to local community standards. The effect of different configurations is examined closely in this report.

While there is no good data on how the different school districts or libraries have decided to configure their filters, an informal survey conducted for this study indicated a wide range in approaches. Accordingly, the study tests a "least restrictive," "intermediate," and "most restrictive" configuration for each filtering product.

Background

Over the past few years, there has been an explosion in Internet use among young people in this country. According to the U.S. Census Bureau, three out of four teenagers (14-17 years old) used the Internet in 2001, up from half of teens in 1998. Overall, the Census Bureau reports that nearly 60% of all school-aged children (5-17) used the Internet in 2001.

As access to the Internet expands, more and more young people are turning to the Web as a source of health information. Teenagers in particular face a host of sensitive health issues—from substance abuse to birth control to eating disorders–and the confidentiality and convenience of the Web make it a popular destination for those seeking information and guidance on those topics. Indeed, earlier studies indicate that more than 70% of 15- to 17-year-olds say they have used the Internet to look up health information. Some of the most common topics searched include sexual health issues such as pregnancy, birth control, HIV/AIDS, or other sexually transmitted diseases (40% have researched one of these subjects); problems with drugs or alcohol (25%); and depression or mental illness (17%).

At the same time, concerns about young people's exposure to online pornography and other adult content led to the passage of the Children's Internet Protection Act (CIPA) in December 2000. CIPA requires schools and libraries receiving federal funds to block access to inappropriate Internet content through the use of "technology protection measures"— i.e., Internet blocking software. The CIPA requirement for *libraries* was struck down in Spring 2002 by a circuit court on the grounds that it violates the First Amendment, although that decision is currently being appealed to the U.S. Supreme Court. Meanwhile, the CIPA requirements for *schools* have *not* been challenged in the courts, and 73% of schools already employ filters of some kind. The law requires blocking of sites that are obscene, contain child pornography, or are "harmful to minors"—a standard it leaves up to the local community to set.

In addition to filters at school, many parents use some sort of filtering product at home as well. In earlier studies, a third of teens with home Internet access have reported that there is a filter in place at their home. While the primary purpose of this study is to review the filters most commonly used in schools or libraries, the most common home product, AOL Parental Controls, was also tested.

Methodology

The study was designed by Kaiser Family Foundation staff in partnership with Dr. Paul Resnick of the School of Information at the University of Michigan and Dr. Caroline Richardson of the University of Michigan Medical School. It was conducted by Drs. Resnick and Richardson along with several of their colleagues. The study used the following methodology:

Researchers simulated young people's searches for health information on 24 health topics, on each of six different search engines that are popular among young people: Yahoo!, Google, America Online (AOL), Microsoft Network (MSN), Ask Jeeves, and Alta Vista.

> The health topics were selected based on three factors: previous national surveys about the subjects most frequently researched online by teenagers, frequency data about commonly used search terms, and researchers' decision to test a variety of subjects including health topics unrelated to sex, topics involving sexual body parts but not related to sex, health topics related to sex, and controversial health topics.

<A> HEALTH TOPICS SEARCHED

Health topics unrelated to sex:
1. Diabetes: Diabetic diet, Diabetes
2. Substance abuse: Ecstasy, Alcohol
3. Mental health: Suicide, Depression

Health topics related to sex:
1. Sexually transmitted diseases: STD, Herpes
2. Safe sex: Safe sex, Condoms
3. Pregnancy prevention: Pregnancy, Birth control

Health topics involving sexual body parts, not sex related:
1. Breast cancer: Breast cancer, cancer
2. Genital infections: Jock itch, Yeast infection
3. Breast feeding: Breast feeding, Breast pump

Controversial health topics:
1. Abortion: Abortion, RU486
2. Homosexuality: Gay, Lesbian
3. Rape: Rape, Date rape

Each unique Web site listed in the first four pages of search results was coded as to whether it was health information, pornography, or "other."

> Health information was defined as information about topics that would be discussed in a medical school or school of public health. Pornography was defined as any text or graphic of a sexual act or genitals designed to appeal to prurient interests that was not of an educational or scientific nature.

The health and pornography Web sites returned from those searches were then tested against a variety of commonly used blocking products.

> The primary purpose of the study is to test filtering products widely used in schools or libraries. Therefore, the results reported here concern six filtering products for institutional use: SmartFilter, 8e6, Websense, CyberPatrol, Symantec, and N2H2. However, researchers also tested one product widely used in the home, AOL Parental Controls, and those results are presented separately in Table G.

Each of the seven blocking products was tested at various levels of configuration.

< B > FILTERING CONFIGURATIONS TESTED

Description	Examples of Categories Blocked
Least restrictive	
Block pornography only, based on the configurations used in another recent test commissioned by the U.S. Department of Justice.	Pornography; Sex acts; Adult/Sexually explicit
Intermediate	
Block categories that may be considered inappropriate; based on the configuration used by one statewide school network.	Drugs; Nudity; Weapons; Hate/discrimination
Most restrictive	
Block any category that might plausibly be blocked in some school or library.	Tobacco; Profanity; Swimsuits; Jokes; Auctions; Games; Dating

> The six products used in schools and libraries all allow the system administrator who installs the product to customize the blocking configuration by selecting topics or categories for blocking. Some products provide one or more default configurations, but calls to schools and libraries revealed a wide variability in their actual configurations.

> For each of the school and library products researchers tested three configurations that were roughly comparable across products: least restrictive (blocking pornography only); intermediate (based on a configuration used by one major statewide school network, which blocks pornography plus several other categories of content that may be considered inappropriate); and most restrictive (set to block all topics or categories that might plausibly be blocked in a school or library).

> Researchers also tested two settings for AOL Parental Controls: "Young Teens" and "Mature Teens," but these settings were not comparable to those tested for the school and library products.

Researchers also simulated a search for pornographic Web sites, using six frequently used search terms, and tested those results against each of the blocking products as well.

Finally, a separate test was conducted on a list of "recommended" health sites, gathered from the teen health sections of the online directories at Yahoo! and Google.
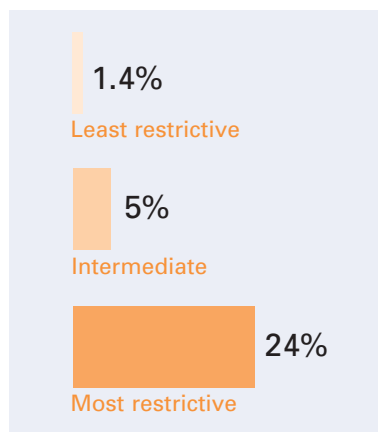
In the end, a total of 3,053 health Web sites and 516 pornography sites were tested against the filtering products (2,467 of the health sites were from the simulated searches, and 586 were "recommended" sites from the online directories).

Key Findings

1. At the least restrictive or intermediate configurations, the filters tested do not block a substantial proportion of general health information sites; however, at the most restrictive configuration, one in four health sites are blocked.

When set at the *least restrictive* level of blocking ("pornography only"), filters block an average of 1.4% of all health sites; this figure is the average across all six of the institutional filters tested, and across all of the various health topics studied.

< C > PROPORTION OF HEALTH SITES BLOCKED BY INTERNET FILTERS, ON AVERAGE, BY LEVEL OF CONFIGURATION

1.4%
Least restrictive

5%
Intermediate

24%
Most restrictive

When set at the *intermediate* level of blocking, filters block an average of 5% of all health sites.

When set at the *most restrictive* configuration, 24% of all non-pornographic health sites are blocked.

2. Some health topics are much more likely to be blocked, even at the least restrictive or intermediate blocking levels.

Even when set at their *least restrictive* blocking configurations, filters block an average of about one in ten non-pornographic health sites resulting from searches on the terms "condoms," "safe sex," and "gay."

At the *intermediate* level of blocking, a substantial proportion of health sites resulting from searches on some sensitive topics are blocked:

Condoms <27% of health sites were blocked>

Ecstasy <25% of health sites were blocked>

Gay <24% of health sites were blocked>

Safe sex <20% of health sites were blocked>

gayhealth.com blocked at least restrictive setting

hivchannel.com blocked at least restrictive setting

3. The proportion of *pornographic* sites blocked *does not* increase substantially based on how the filters are configured.

At the *least restrictive* setting, the filtering products block an average of 87% of all pornographic sites; at the *intermediate* level, an average of 90% of porn sites are blocked, and at the *most restrictive* configuration, 91% of porn sites are blocked.

4. The proportion of *non-pornographic* health sites blocked *does* increase

substantially based on how the filters are configured, especially on topics related to sexual health.

As noted, across all health topics studied, the proportion of non-pornographic sites blocked averaged 1.4% at the *least restrictive* blocking level, 5% at the *intermediate* blocking level, and 24% at the *most restrictive* blocking level. Similar results emerge for the "recommended" health sites listed on the teen health sections of the online directories at

Yahoo! and Google. At the *least restrictive* setting, an average of 0.5% of these sites (on a variety of health topics) were blocked, at the *intermediate* setting an average of 2.8% were blocked, and at the *most restrictive* setting an average of 24% were blocked.

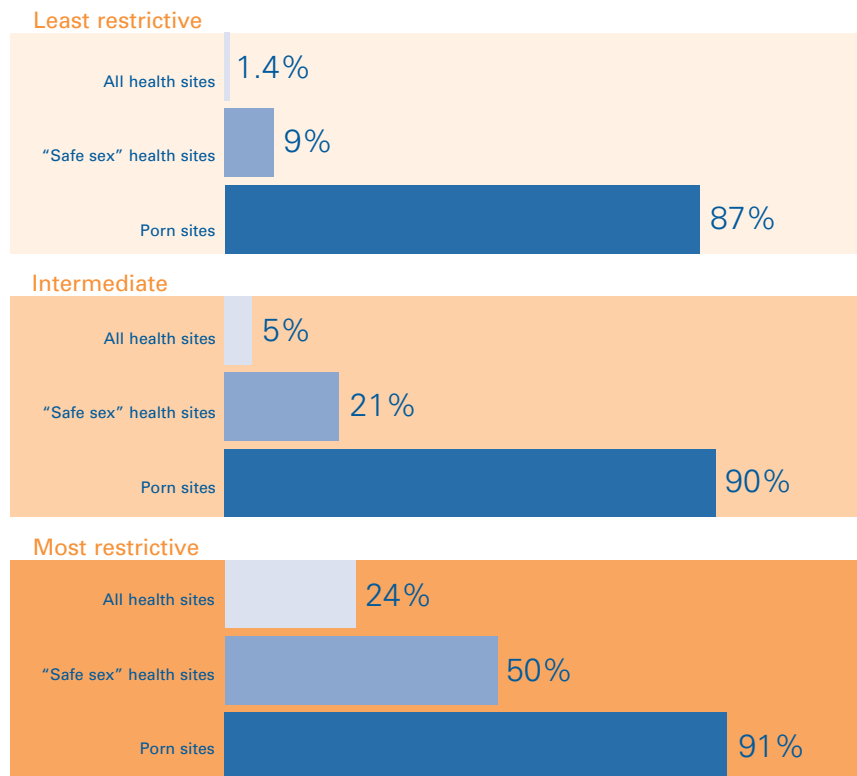The increase in blocked health content is especially pronounced on searches related to *sexual health*. For example, for a search on "safe sex," on average about one in ten health sites (9%) is blocked at the *least restrictive* level of blocking, one in five (21%) at the *intermediate* level, and one in two (50%) at the *most restrictive* level.

5. Incidental exposure to pornography during health information searches does not appear to be a substantial problem; filters can further reduce but not eliminate such incidental exposure.

Across the 24 health-related searches conducted on six different search

Key Findings Continued

<D> PROPORTION OF PORNOGRAPHIC AND HEALTH SITES
BLOCKED, ON AVERAGE, BY TOPIC AND LEVEL OF BLOCKING

Least restrictive

All health sites	1.4%
"Safe sex" health sites	9%
Porn sites	87%

Intermediate

All health sites	5%
"Safe sex" health sites	21%
Porn sites	90%

Most restrictive

All health sites	24%
"Safe sex" health sites	50%
Porn sites	91%

At the *least restrictive* setting, the percentage of health sites blocked by the different filters ranged from 0.6% to 2.3%; at the *intermediate* level the range between products was 2.8% to 7.6%; and at the *most restrictive* level the range was 15.1% to 35.4% (see Table G).

7. Many Web sites of health organizations that provide online content for youth are being blocked by one or more filtering products.

In the previous findings, we have been examining the filters' performance by looking at the average rate of over- or under-blocking for any one of the six products studied. For example, when set at the intermediate level, on average any one filter blocks 5% of health sites. However, cumulatively, the six filters in this study blocked a total of 16% of all health sites at the intermediate level. Therefore, 16% of all health sites would be blocked by one

engines, 1% of the results contained pornographic content.

At the minimal blocking configurations, the products blocked an average of 62% of these "inadvertently" retrieved porn sites. Thus, the products were noticeably worse at identifying pornography

resulting from health searches than from deliberate searches for pornography (87% blocking rate).

6. The different filtering products vary somewhat in the proportion of health sites they blocked, especially at the more restrictive settings.
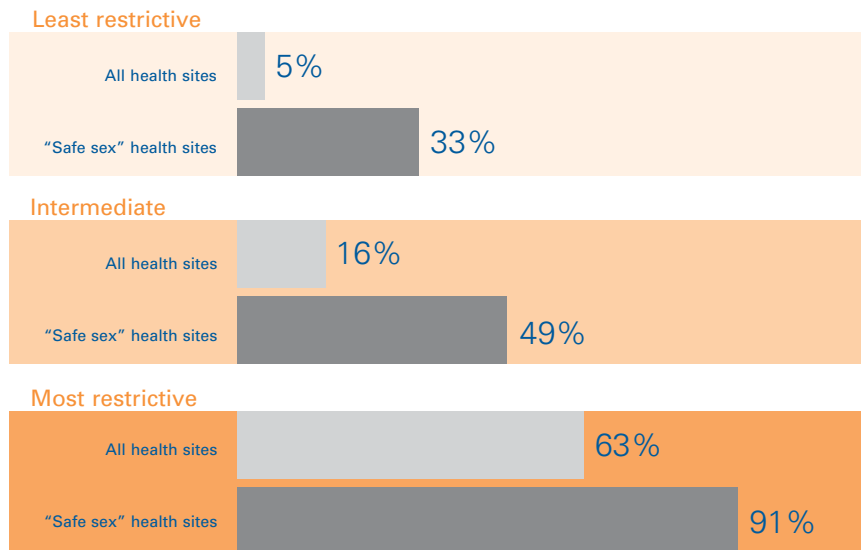
Organizations providing information on *sexual health* are even more likely to have their Web sites blocked by some filtering product. For example, one in every three of the safe sex health sites studied (33%) were blocked by at least one filtering product at the *least restrictive* setting, one in two (49%) were blocked by at least one filter at the *intermediate* level, and more than nine in ten (91%) were blocked at the *most restrictive* setting.

ripnroll.com blocked at least restrictive setting

filtering product or another, when set at the intermediate level.

Across all of the health topics studied, 5% of all health sites were blocked by at least one of the filtering products studied when set at the *least restrictive* configuration, 16% were blocked by at least one filtering product at the *intermediate* configuration, and 63% were blocked by at least one filter at the *most restrictive* configuration.

<E> PROPORTION OF HEALTH SITES BLOCKED BY ONE OR MORE FILTERING PRODUCT, BY LEVEL OF CONFIGURATION AND BY TOPIC

Least restrictive

All health sites	5%
"Safe sex" health sites	33%

Intermediate

All health sites	16%
"Safe sex" health sites	49%

Most restrictive

All health sites	63%
"Safe sex" health sites	91%

Key Findings Continued

8. Internet users seeking health information can avoid some blocking by using alternate search terms.

Sexual health sites associated with certain search terms—such as "safe sex" or "condoms"—were much more likely to be blocked than sites that resulted from alternate search strings, such as "birth control," "STD," or "herpes" (for example, 28% of health sites from the "condom" search were blocked at intermediate blocking levels, compared to 5% of sites from a search on "birth control").

< F > BLOCKING, BY ALL PRODUCTS, ON HEALTH SEARCHES

Search String	Sites Returned From Searches			Health Information Sites Blocked (%)		
	Health sites	Porn sites	Other sites	Least*	Intermediate*	Most*
Category 1 < Health topics unrelated to sex >						
Diabetic diet	202	0	28	0.1	0.2	13.9
Diabetes	196	0	34	0.1	0.4	10.0
Ecstasy	171	2	64	0.3	24.9	36.2
Alcohol	204	0	27	0	7.1	12.7
Suicide	200	2	32	0.2	1.7	13.7
Depression	189	0	46	0	1.0	11.2
Category 2 < Health topics involving sexual body parts, not sex related >						
Breast cancer	224	0	8	0	0.2	6.9
Cancer	221	0	12	0	0.3	3.7
Jock itch	180	0	11	0.6	1.4	15.4
Yeast infection	206	1	11	0.1	1.1	18.4
Breast feeding	217	0	18	0.2	1.2	18.6
Breast pump	163	0	69	0	0.9	26.3
Category 3 < Health topics related to sex >						
STD	112	0	118	1.4	3.4	23.4
Herpes	203	0	29	1.0	1.8	23.0
Safe sex	131	11	51	9.3	20.5	50.0
Condoms	152	9	72	9.1	27.7	55.4
Pregnancy	219	0	15	0.4	0.6	31.6
Birth Control	210	0	27	1.8	5.0	34.7
Category 4 < Controversial health topics >						
RU486	177	1	54	0.3	2.1	25.8
Abortion	184	0	48	0.2	3.2	44.6
Gay	74	10	113	11.1	24.6	59.9
Lesbian	79	7	149	3.8	17.1	59.0
Rape	151	2	74	1.2	3.3	22.0
Date Rape	152	1	30	1.7	5.2	21.1

* The blocking percentages reflect the percentage of unique sites returned by all searches on that search term that were blocked. If a site is suggested by more than one search engine, it is counted only once. Percentages are averaged across the six products for the given setting (least restrictive, intermediate, or most restrictive).

\<G\> BLOCKING RESULTS FOR ALL PRODUCTS ACROSS THREE LEVELS OF BLOCKING RESTRICTIVENESS

	SmartFilter	8e6	Websense	CyberPatrol	Symantec	N2H2	AVERAGE*	AOL PC
Percent of Health Information URLs Blocked (n=2,467)								
	2.3	1.1	0.6	1.6	1.9	0.8	1.4	
	5.8	4.5	3.8	2.8	7.6	6.5	5.2	3.2
	18.2	15.1	35.4	22.4	33.5	19.5	24.0	16.1
Percent of Pornography URLs Blocked (n=516)								
	87.2	89.1	83.9	85.7	87.8	89.5	87.2	
	88.7	90.9	91.3	85.7	89.3	92.8	89.8	92.1
	89.0	92.1	93.8	87.2	90.5	94.0	91.1	94.8
Percent of Recommended Health Information URLs Blocked (n=586)								
	0.0	0.5	0.5	0.5	1.4	0.3	0.5	
	1.0	1.2	1.4	0.9	8.4	3.8	2.8	1.2
	16.8	10.9	39.4	26.5	28.5	23.2	24.2	15.0

*Average of all products, excluding AOL PC

▢ Least restrictive   ▢ Intermediate   ▢ Most restrictive

\<H\> PERCENT OF NON-PORNOGRAPHIC HEALTH SITES ON "CONDOM,""GAY," AND "SAFE SEX" THAT WERE BLOCKED, AT THE LEAST RESTRICTIVE LEVEL, BY PRODUCT (N=214)

SmartFilter	CyberPatrol	Symantec	8e6	N2H2	Websense
16%	15%	10%	8%	5%	4%

Conclusion

The extent to which Internet filters may adversely impact young people's access to online health information depends in large measure on how the filters are configured by the schools or libraries installing them. Indeed, how a specific filtering product is configured is more important in avoiding inadvertent blocking of health information than is the choice between different products.

The major filtering products on the market allow administrators a great deal of leeway in determining how much material will be blocked. At their least restrictive settings, overblocking by filters has a negligible impact on access to general health information—especially when compared to other factors that can affect search results, such as spelling errors, limited search skills, and uneven quality of search engines. However, even at their least restrictive settings, filters could

have a modest impact on those seeking information on sexual health issues; on average, filters incorrectly blocked about one in ten sites on safe sex, condoms, or health issues pertaining to gays. A determined searcher would likely still find the information he or she was searching for, particularly by altering search terms slightly: far fewer health sites were improperly blocked on terms such as "birth control," "STD," or "herpes."

The use of Internet filters could have a major impact on access to health information if administrators at schools and libraries configure them to block additional categories of content beyond just pornography: one in every four health sites was blocked, on average, at the most restrictive set-tings, and for topics related to sexual health, the rates were as high as one in every two sites.

While using a more restrictive setting for the filters results in a significant

increase in blocking of health sites, it yields only a marginal increase in effectiveness at blocking pornography (the percentage of health sites blocked increases from 1.4% to 24%, while the percent of porn sites blocked increases from 87% to 91%).

Inadvertent exposure to pornography occurs relatively infrequently during health searches—just 1% of the time—and none of the filters tested was particularly effective at blocking these sites (62% were blocked, on average).

Young people who *proactively* seek out pornographic Web sites will find the filters more of an obstacle, but those who are determined to find such content will be able to do so eventually, even at the most restrictive levels of blocking: if nine out of ten porn sites are blocked, there will still be those students who pursue and find the one in ten that gets through.

Organizations that provide online health information for adolescents need to be aware of the potential impact of Internet filters on their efforts, given that nearly two out of three health sites are blocked by at least one of the filtering products when set at their most restrictive levels. For those providing sexual health information, the issue is quite serious: one in three "safe sex" health sites are blocked by at least one of the filters even when set at their *least* restrictive settings.

In sum, the results of this study indicate that Internet filters reduce but do not prevent inadvertent exposure to pornography during searches for health information, something that is an infrequent occurrence to begin with. They make it substantially harder for, but do not entirely prevent, young people from *proactively* seeking out pornographic content.

The filters do not significantly impede searches for general health information when they are set at low levels of blocking, but they do have a modest impact on searches for sexual health information even at these settings. Filters do interfere with general health searches at higher levels of configuration, and have an especially serious impact on searches for *sexual health* information at these levels, blocking many non-pornographic sites.

Setting Internet filters to more restrictive levels does not significantly increase their effectiveness at blocking pornography, but does significantly increase overblocking of non-pornographic health content. Some schools or libraries may decide that they wish to block additional content categories beyond pornography, but this should be a conscious decision by policymakers rather than an

inadvertent one by those who install the filter software; and this decision should be made with an awareness of the impact on young people's access to health information.

13