SUPPLEMENTAL REBUTTAL EXPERT REPORT OF
PHILIP B. STARK, PH.D.

1 AUGUST 2006

ACLU v. Gonzales

Civ. Action No. 98-5591 (E.D. Pa.)

SUPPLEMENTAL REBUTTAL EXPERT REPORT OF
PHILIP B. STARK, PH.D., 1 AUGUST 2006

1. This report estimates the effectiveness of two additional content filters[1] to

address claims in the 6 July 2006 Rebuttal Expert Report and the 26 July 2006

Supplemental Expert Report of J. Christopher Racich. The estimates rely on a

database provided by CRA International.

2. CRA International tested the filters Mr. Racich used; the settings and testing

protocol will be described in the 1 August 2006 Supplemental Rebuttal Report of

Paul Mewett. I used the CRA International results to estimate the underblocking

rate (failure to block sexually explicit webpages[2]) and overblocking rate (erroneous

blocking of clean webpages[3]) of each filter on various populations of webpages. I

calculated confidence limits for some of the estimates.[4]

3. Table 1 gives estimates of underblocking and overblocking and Table 2 gives

lower confidence limits, for the Google and MSN indexes. Table 3 gives estimates

of the percentage of domestic sexually explicit webpages among the sexually

explicit webpages in the Google and MSN indexes that filters do not block. Table

---

1   8e6 and SafeEyes.
2   A "sexually explicit" webpage is one that CRA International would put in category 5f.
3   A "clean" webpage is one that CRA International would put in category 1a.
4   The populations of webpages and the methodology are described in my 8 May 2006
    expert report.

4 gives estimates of overblocking and underblocking for webpages retrieved by AOL, MSN and Yahoo! queries. Table 5 gives lower confidence limits for the percentage of AOL, MSN and Yahoo! queries that return at least one sexually explicit webpage that would not be blocked by filters. Table 6 shows underblocking and overblocking for Wordtracker queries.   The filters Mr. Racich used perform comparably to filters CRA International had tested previously.

Filter[5]	Underblocking		Overblocking	
	Google	MSN	Google	MSN
1	8.9%	8.6%	22.6%	23.6%
2a	16.8%	18.7%	19.6%	10.3%
2b	17.7%	20.5%	21.9%	18.9%
3a	38.3%	45.4%	2.8%	3.0%
3b	28.3%	46.7%	1.4%	0.7%
4	31.0%	33.5%	1.4%	0.9%
5a	12.7%	16.5%	3.6%	4.1%
5b	12.4%	18.9%	4.0%	3.7%
6	16.1%	26.0%	12.4%	13.2%
7	44.0%	46.1%	3.3%	2.2%
8a	60.2%	54.9%	1.4%	0.7%
8b	58.4%	54.2%	0.9%	0.4%
9	41.8%	40.3%	9.4%	5.7%
10	18.3%	23.0%	9.4%	7.5%
11	16.2%	15.2%	3.3%	3.2%

Table 1: Estimated underblocking and overblocking of webpages in the Google and MSN indexes. Among sexually explicit webpages, the percentage that are not blocked by a filter is the rate of underblocking. Among clean webpages, the percentage that are blocked by a filter is the rate of overblocking.

---

5   The filters are as follows; settings are described more fully in the 8 May 2006 Expert Report of Paul Mewett, the 6 July 2006 Rebuttal Report of Paul Mewett, and the 1 August 2006 Supplemental Rebuttal Report of Paul Mewett. 1: AOL Mature Teen. 2a: MSN Pornography. 2b: MSN Teen. 3a: ContentProtect Default setting. 3b: ContentProtect Custom setting. 4: CyberPatrol Custom setting. 5a: CyberSitter Default setting. 5b: CyberSitter Custom setting. 6: McAfee Young Teen. 7: Net Nanny Level 2. 8a: Norton Default setting. 8b: Norton Custom setting. 9: Verizon. 10: 8e6. 11: SafeEyes.

Filter	Underblocking		Overblocking	
	Google	MSN	Google	MSN
1	5.6%	6.5%	18.4%	21.0%
2a	12.1%	15.7%	15.8%	8.5%
2b	12.8%	17.4%	17.8%	16.6%
3a	31.3%	41.3%	1.5%	2.1%
3b	22.2%	42.6%	0.6%	0.4%
4	24.6%	29.7%	0.6%	0.5%
5a	8.6%	13.6%	2.1%	3.1%
5b	8.4%	15.9%	2.4%	2.7%
6	11.4%	22.5%	9.3%	11.3%
7	36.8%	41.9%	1.9%	1.5%
8a	52.9%	50.7%	0.6%	0.4%
8b	51.1%	50.1%	0.4%	0.2%
9	34.7%	36.2%	6.7%	4.4%
10	13.1%	19.6%	6.7%	6.0%
11	11.4%	12.3%	1.9%	2.3%

Table 2: 95% lower confidence limits for the entries in Table 1. For illustration, at 95% confidence, filter 2b fails to block at least 12.8% of the sexually explicit webpages in the Google index. Similarly, at 95% confidence, filter 2b blocks at least 16.6% of the clean webpages in the MSN index.

Filter	Estimated Domestic Underblocking	
	Google	MSN
1	40.0%	40.6%
2a	31.6%	42.9%
2b	40.0%	37.7%
3a	39.0%	45.8%
3b	40.6%	47.1%
4	48.6%	44.0%
5a	50.0%	32.8%
5b	57.1%	36.2%
6	44.4%	37.5%
7	41.7%	48.1%
8a	35.3%	49.3%
8b	36.4%	49.7%
9	37.0%	42.4%
10	42.1%	46.8%
11	35.3%	40.4%

Table 3: Of the sexually explicit webpages in the Google and MSN indexes that filters do not block, the estimated percentage that are domestic webpages.

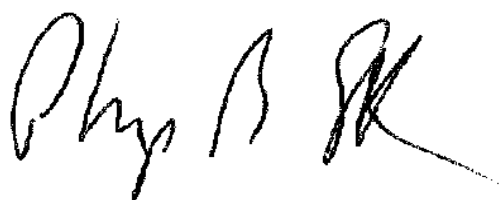| Filter	Underblocking for results	Overblocking for results	Domestic Underblocking	Underblocking for queries
1	6.2%	12.5%	57.0%	15.6%
2a	21.4%	4.4%	86.1%	32.3%
2b	20.8%	5.8%	91.9%	28.1%
3a	18.4%	6.4%	70.1%	46.2%
3b	20.4%	0.0%	62.1%	42.2%
4	34.6%	0.4%	94.9%	65.6%
5a	11.2%	4.6%	33.8%	23.2%
5b	10.0%	5.3%	44.1%	20.1%
6	14.2%	20.7%	80.7%	30.9%
7	28.1%	3.7%	79.4%	36.6%
8a	42.1%	0.8%	85.3%	51.6%
8b	43.4%	0.0%	85.6%	56.1%
9	23.1%	1.3%	80.9%	41.6%
10	7.3%	7.5%	78.0%	23.4%
11	13.7%	1.9%	87.8%	29.8%

Table 4: Estimated underblocking and overblocking of the results of AOL, MSN and Yahoo! searches. "Underblocking for results" is the percentage of sexually explicit search results that are not blocked. "Overblocking for results" is the percentage of clean search results that are blocked. "Domestic underblocking" is the percentage of domestic webpages among sexually explicit webpages the filters do not block. "Underblocking for queries" is, among queries that retrieve any sexually explicit webpages, the percentage that retrieve at least one sexually explicit webpage that is not blocked.

Filter	Underblocking for queries
1	5.3%
2a	20.9%
2b	18.8%
3a	10.0%
3b	25.4%
4	24.4%
5a	11.2%
5b	8.1%
6	10.4%
7	20.8%
8a	49.3%
8b	54.3%
9	31.4%
10	11.7%
11	14.9%

Table 5: Lower 95% confidence limits for the rightmost column in Table 4.

Filter	Underblocking for results	Overblocking for results	Domestic Underblocking	Underblocking for queries
1	1.3%	19.6%	69.2%	4.3%
2a	2.7%	13.3%	86.1%	8.2%
2b	2.6%	13.7%	83.1%	8.3%
3a	7.5%	12.4%	84.1%	23.1%
3b	8.1%	7.8%	84.9%	25.3%
4	3.9%	9.2%	86.4%	10.1%
5a	1.4%	19.9%	69.3%	5.1%
5b	2.9%	18.2%	84.0%	9.4%
6	2.8%	32.8%	70.7%	9.3%
7	12.6%	9.5%	82.9%	34.4%
8a	9.9%	4.8%	79.4%	25.2%
8b	10.2%	2.9%	79.4%	25.9%
9	4.4%	16.1%	67.9%	15.0%
10	3.4%	25.1%	93.0%	10.3%
11	2.0%	16.5%	96.6%	6.4%

Table 6: Underblocking and estimated overblocking for the results of Wordtracker queries. "Underblocking for results" is the percentage of sexually explicit search results that are not blocked. "Overblocking for results" is the percentage of clean search results the filter blocks. "Domestic underblocking" is the percentage of domestic webpages among the sexually explicit webpages the filters do not block. "Underblocking for queries" is, among the queries that retrieve any sexually explicit webpages, the percentage that retrieve at least one sexually explicit webpage that is not blocked. Overblocking was estimated from a random sample of clean search results. Underblocking was determined from all the sexually explicit search results.

_____ Dated _1_ August 2006.

Philip B. Stark