Sheffield Hallam University

Investigating the unofficial factors in Google ranking

MARDANI, Amir, AKHGAR, Babak <http://orcid.org/0000-0003-3684-6481>, ANDREWS, Simon <http://orcid.org/0000-0003-2094-7456>, YATES, Simeon and HASSANZADEH, Mohammad

Investigating the Unofficial Factors in Google Ranking

Amir Mardani, Babak Akhgar, Simon Andrews, Simeon Yates [1,] and Mohammad Hassanzadeh[2]
1.Faculty of ACES, Sheffield Hallam University, Sheffield, UK
2.Tarbiat Modares University, Tehran, IRAN
inadram@gmail.com, s.andrews@shu.ac.uk, b.akhgar@shu.ac.uk, hasanzadeh@modares.ac.ir

Abstract

This paper evaluates the effectiveness of some "unofficial" factors in Search Engine Optimisation. A summary of official Google guidelines is given followed by a review of "unofficial" ranking factors as reported by a number of experts in the field of Search Engine Optimisation". These opinions vary and do not always agree. Experiments on keyword density, web page titles and the use of outbound links were conducted to investigate the expert's hypotheses by analysing Google result pages. The results demonstrate that webmasters should avoid having unnecessary outbound links, while attempting to repeat the important keywords of each page one time in their titles, to increase the pages ranking in the results page.

Keywords: SEO, Search Engine Optimisation, SEO unofficial factors

1. Introduction

Every month, more than eighteen billion web searches are performed on the Internet [1]. For companies and individuals have become reliant on the "lower cost", "focus" and simplicity of the Web as a route to market, customers and clients [2] [3]. Therefore sagacious business managers, looking for ways to improve their website's ranking status in Search Engines Result Pages (SERPs) use Search Engine Optimisation methods (SEOs).

In addition to well-known SEOs, based on published factors in Google's ranking process, there are a number of *unofficial* ranking factors that have never been confirmed or denied by Google, that SEOs may exploit. This paper investigates some of these unofficial factors and explores some of the variables involved to thereby recommend appropriate SEOs to exploit them.

Firstly, SEO is explained in relation to official ranking factors published by Google [4]. The research then focuses on unofficial factors which may have an effect on the ranking of a website in search results. The outcomes of this research could be useful for webmasters and site owners who want to augment their viewer density through the Google search engine.

2. A review of SEO factors

There are numerous search engines but only some of them have been successful in attracting large numbers of users [5]. It therefore makes sense for webmasters to implement SEOs that target the most widely used search engines. This segment of the study examines SEO factors pertinent to Google, arguably the most important search engine [6, p. 1].

What is SEO?

SEO is set of small modifications to segments of a website that can assist in getting more hits from search engines [7, p. 1]. There are over two hundred signals that Google considers when ranking websites while scoring their respective search result [8] but Google, in a guidelines for webmasters, officially only cites a limited number of them.

Official factors

Google introduced useful tactics and factors that can help webmasters get a better accessibility status in Search Engine Results Pages (SERP). Google has guidelines relating to page title, site speed, content, anchor text, URLs, navigation, head tags, images and links. Although following these guidelines is certainly effective and can assist search engines to index and crawl websites more easily, they "won't tell you any secrets that'll automatically rank your site first for queries in Google (sorry!)" [4]. A summary of each guideline follows:

Title

Google suggests that webmasters should have unique titles which describe the content of each page accurately [4].

Site speed

"Site speed shows how quickly a website responds to web requests" [9]. Google includes this signal in its search ranking algorithm to encourage webmasters to compact their website [10] .

Content

Creating unique and fresh content for users with relevant information helps Google to reach its goal to "give people the most relevant answers to their queries

as quickly as possible" [11]. Therefore, useful content is one of the most important signals that Google considers in its ranking algorithm. Google uses various criteria to evaluate the quality of the content such as checking the similarity of the content, attractiveness of the topic for the visitors, rationality and comprehensiveness [12].

URL
Google considers the URL of the pages as a signal for ranking websites [13] and asks webmasters to have a descriptive URL for categories and filenames [4].

Navigation
Navigation can help Google to find out important content of each website as well as guiding visitors to find their desired content quickly. Google suggests webmasters plan navigation based on their homepage wisely with a "navigational menu", "text-based links" or a "user-viewable site map" [14].

Anchor text
Anchor text is a clickable text that a user sees on a link [4]. Google asks webmasters to have short but descriptive anchor texts to describe the content and importance of their pages to search engines [15] [4].

Head tags
Webmasters can use concise phrases when describing the content of a page via multiple HTML heading size tags such as "<h1>", "<h2>" and "<h3>". These are important to inform the search engine about the hierarchical structure of the website and the relative importance of text. Although styling the text might achieve the same visual presentation, it does not provide the same meaning or metric to the search engine that a head tag does [8].

Optimise images
Google suggests webmasters put related content around their images and use brief but descriptive text in the "alt" attribute to provide image-related information for their pages. In addition it is quite useful to have a brief but descriptive file name for images rather than generic names such as "pic.gif" or "1.jpg". Google also asks that images be grouped according to size into directories [4] [16] [17] to help Googlebots distinguish the topic of their pages [8].

Link
A website with a proper linking structure can help both Google and users to have better exploration experience and also help it to achieve better visibility in search results [18]. Google uses mature text-matching algorithms to return pages which are both relevant and important for each search query and links are one of the most important factors which can get pages "authority" and "importance". In fact, Google consider a link between pages A to B as a *vote* from A to B and the importance of page A is carried over to page B as "link juice".

On the other hand, Google penalises websites which try to manipulate the search engine by putting unnecessary keywords in their content or copyright content at their end [19]. Google strictly asks webmasters to avoid using keywords excessively in their URLs, Anchor text and images [4]

Google Unofficial factors
Although aligning the website structure and functionality with official factors is good practice, using effective unofficial factors can act as a powerful competitive advantage. Unofficial ranking factors are extensively argued over by SEO experts. Some of these factors are rejected by search engines as cheating, such as "link farming" [20], "clock threading" [21], "hidden text" [22] and "automated queries" [23] but there are other methods that may be effective that are neither officially accepted or rejected by Google. The following sections examine some of these unofficial SEO factors, namely "Best title", "Keyword density" and "outbound link".

Best title
"Do keep it short" says Grappone and Couzin [24, p. 173]. Most search engines present only the first 60 characters of the title in their search result; therefore webmasters should keep their titles short [7, p. 64] [24, p. 173] [25, p. 60] [26, p. 29]. In addition Grappone and Couzin strongly recommended avoiding repeating keywords in titles [24, p. 173]. Similarly Peter Kent believes in short titles but recommends inclusion in the title of the most important keyword of the page [27, p. 35]. However, Konia in "WebPosition Gold", a famous "black hat" SEO tool, recommends webmasters use their primary keywords in the title tag *at least* once. He said webmasters can attract more traffic by using the same keyword in the title multiple times but in different rows. He also stood against the short title idea and suggests webmasters can use longer titles to achieve a better position in search results [28, p. 133]. Enge *et al.* also advocate long titles: "Target longer phrases if they are relevant" [6, p. 212]. Enge *et al.* and Fox believe that having more accurate and descriptive titles are better than simple titles which may be ambiguous or convey less information about the content [6, p. 212] [29, p. 147]. However Google suggests both views have merit, recommending titles that are brief but also descriptive [4].

Keyword density

Keyword density, or in other words the number of times that a specific keyword is repeated in the content, is one of the most important factors that almost all SEO experts believe in. However, there are different points of view about the best keyword density percentage for generating better results.

Jerkovic believes that a good keyword density is between 0.2% and 4%. At the same time he claims that if you go beyond 10%, search engines will penalise you [7, p. 67]. Also, the vendor of WebPosition Gold argues that this percentage could vary from 1% to 4% according to your targeted search engine [28, p. 19]. On the other hand, Kent [27, p. 105] and Baylin [26, p. 135] do not believe that keyword density is a major factor at all. Similarly, Enge *et al*. believe that search engines use more sophisticated analyses than simply counting keywords [6, p. 158]. However, although Google does not encourage webmasters to repeat their keywords within the content of their websites, it has never denied the role of keyword density in SERP.

Outbound links

Outbound links refer to the links which point to external websites. There are webmasters that worry about making outbound links because they think it might cause them to lose their PageRank and also their visitors when they are sending them out of their website. On the other hand, there are some who believe that having only inbound links with no outbound links limits the scope of their website and reduces the quality and richness of the user's experience, and that the best plan is to have a balance between the two [30, p. 268]. Linking to other sites might at first seem ill-advised, in that visitors are being directed away, but it can help visitors find relevant sources. Search engines will find out that you are adding value to the web and improve your site's ranking as a consequence [31, p. 43], particularly when there are links to well-known websites [26, p. 160]. Peter Kent also believes that having good outbound links can help [27, p. 430] while Jerkovic states that having outbound links can actually reduce a website's popularity regardless of the quality of the links. High quality target pages could be considered those having high relevancy or are themselves ranked high. [7, p. 92]. Enge *et al*. believe that having outbound links to mistrusted or poor quality websites can hurt a website's reputation and it's ranking [6, p. 52]. Engaging in so-called linking schemes, where co-operative interlinking of websites is encouraged in an attempt to boost ranking, can back-fire and end up having a negative effect on the ranking.

3. Research Methods

To determine the effect of variables involved in the unofficial SEO methods, an empirical case study on various pre-defined websites was carried out, in a controlled experimental environment. To be confident that rankings were only being affected by the variables under investigation it was important that the other factors were the same in all of the websites. These control factors, such as link structure, site speed and content were ensured by using commercial SEO tools such as *opensiteexplorer.org* and *webseoanalytics.com*.

Population and sample

Data collection came from seocasestudy.co.uk sub-domains which have suitable features to control unwanted factors in SEO experiments. *Seocasestudy.co.uk* is a fresh domain that uses HTML pages for testing SEO approaches in a controlled experimental environment.

Results were collected with Google Custom Search (CSE) [32]. CSE uses the same technology that *Google.com* has and takes into account all the factors which *Google.com* cares about [33].

4. Experimental Findings.

Best Title Experiment

Titles in this experiment are varying in length and keyword repetitions. Three word phrases were created, such as "Top love songs", from which to devise page titles and search terms. Search terms were used consisting of one, two or all three words from the phrase. Page titles were created, given in Table 1, using words X, Y and Z, where X was the first word in the phrase, Y the second and Z the third.

Table 1: Results of the Best Title experiment

ID	Title	Average Rank
1	XYZ	4.8
2	XY	6.9
3	X	8.3
4	XYZ | X	5.4
5	XYZ | XY	4.5
6	XYZ | XYZ	4.5
7	XYZ and XYZ	2.2
8	XYZ | XYZ | XYZ	3.7
9	XYZ | XYZ | XYZ| XYZ	5.2

These nine combinations were tested with various phrases, several times. The same content with the same link structure, and keyword density were published in *seocasestudy.co.uk* subdomains to remove any unwanted factors that might effect the results.

The findings given in Table 1 support Enge *et al*. and Fox, they indicate that having a long title does not harm the rank of web pages. For instance, the titles of the pages in category seven are long in comparison with the search term but still have the best position in the SERP [6, p. 212] [29, p. 147]. The findings also provide evidence that pages that do not have all the search term's keywords in their titles rank lower. High ranks are achieved by having each keyword appear in the title at least one time. The results contradict Grappone and Couzin who argue against having duplicate keywords in the titles [24, p. 173] but support Konia and Kent's idea to repeat the keywords in titles [27, p. 35][28, p. 133]. It also seems that connecting the keywords in the title in a meaningful way could be quite useful. For instance, category seven ranks better than category six by using "and" to give a more meaningful title.

Keyword Density Experiment

This experiment sought to find the best keyword density to rank better in Google Search engine result page. The experiment was repeated several times for twenty different densities and search terms consisting of one, two or three keywords. For each experiment, the same content and link structure were published in *seocasestudy.co.uk* subdomains to remove any unwanted factors that might effect the results.

Table 2: Results of the Keyword Density experiment

ID	Keyword density			Comparative Rank
	3 word	2 word	1 word	
1	0.0%	0.0%	0.0%	20
2	1.1%	0.8%	0.4%	19
3	2.3%	1.6%	0.8%	18
4	3.4%	2.3%	1.1%	17
5	4.5%	3.0%	1.5%	16
6	5.5%	3.7%	1.8%	15
7	6.4%	4.3%	2.1%	14
8	7.4%	4.9%	2.5%	13
9	8.3%	5.6%	2.8%	11
10	9.3%	6.2%	3.1%	9
11	10.2%	6.8%	3.4%	10
12	10.9%	7.3%	3.6%	8
13	11.7%	7.8%	3.9%	7
14	12.4%	8.3%	4.1%	6
15	13.1%	8.8%	4.4%	5
16	13.9%	9.3%	4.6%	3
17	14.6%	9.7%	4.9%	2
18	15.2%	10.2%	5.1%	1
19	15.9%	10.6%	5.3%	4
20	16.3%	10.9%	5.4%	12

The findings, given in Table 2, do not support the hypotheses put forward within the literature. Jerkovic believes that Google will penalise pages whose keyword densities go beyond 10%. However, using the data from Table 2, a Hässe diagram (Figure 1) was created. In a Hässe diagram, objects (unshaded boxes) are associated with attributes (shaded boxes) that can be reached by traversing upwards from the object. By scaling keyword density against rank, the diagram indicates that the top four out of 20 ranks (IDs 16, 17, 18, 19) had keyword densities between 12 and 18% for a three word search term, between 8 and 12% for a two word search term and between 4 and 6% for a one word search term. The lowest three rankings (IDs 1, 2, 3) had keyword densities of less than 3, 2 and 1%, respectively, for three word, two word and one word search terms; in other words, the higher the density, the higher the ranking. The results also found a linear relationship between keyword density and number of words in the search term. The density ranges used in the scaling were created proportionally to the density values in each successive group of four rankings (ranks 1-4, 4-8, 9-12, 13-16, and 17-20). This revealed a strong proportional relationship between keyword density and number of search term words. This suggests that variation in the number of search term words is not significant in determining ranking and that the function of increase in ranking by increasing keyword density is linear.

Figure 1: Hässe diagram of ranking against keyword density for one, two and three word search terms
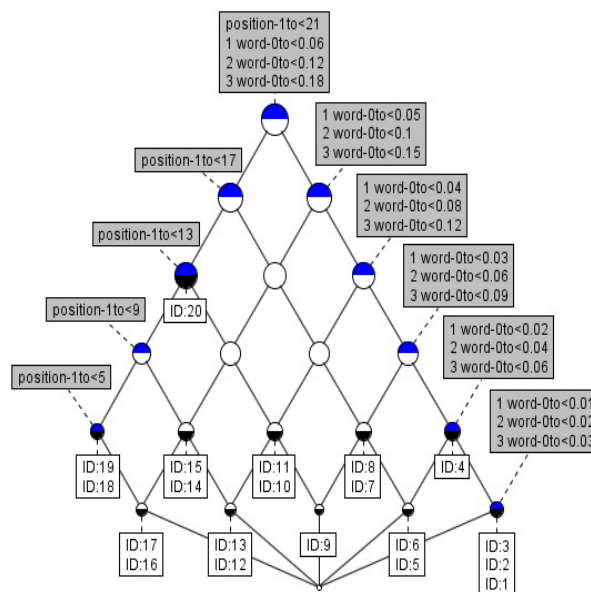
Outbound Link Experiment

The experiment sought to find out if pages that have outbound links to high quality content rank better compared to ones which link to low quality content, or have no outbound links at all. In each experiment, the same content and link structure, page title and keyword density were published in *seocasestudy.co.uk* subdomains to remove any unwanted factors that might effect the results.

Table 3 presents the results of the experiment in 9 different groups of pages. The pages in each group were created with the same values, where PR represents the page rank of the target pages and Description is the description of the type of outbound link used to the target page. The PR varies from N/A (has no ranking at all) to 5. To be more precise, pages which are placed in group one have no outbound links at all. Pages in groups 2 to 7 have similar anchor text to the search term whereas pages in group 8 had dissimilar anchor text. Pages in group 9 had indirect outbound links which means that users go through an intermediate page to reach the target page.

Table 3: Results of the Outbound Links experiment

ID	PR	Description	Average Rank
1	n/a	No outbound link	1
2	3	Similar text	6.5
3	n/a	Similar text	6.4
4	4	Similar text	6.4
5	4	Similar text	6.1
6	2	Similar text	6.4
7	5	Similar text	7
8	4	Dissimilar text	2.9
9	4	Indirect text	2.1

Pages which had no or indirect outbound links ranked better in comparison with other pages. The results support Jerkovic's theory that having outbound links reduces the popularity of a webpage regardless of their quality [7, p. 92]. The experiment did not find any strong correlation between having high quality outbound links and getting a better position in SERP.

On the other hand, it seems that pages whose outbound links have similar anchor text to the search term are ranked lower than those with different anchor text. In other words, in searching for "Y", pages which use "X" for their anchor text rank better in comparison with the ones which link to the same page by "Y" anchor text.

5. Conclusion

This paper evaluates some of the unofficial Google website ranking factors put forward by a number of respected SEO experts.

Research findings indicate that titles of the pages should contain search terms at least one time and at the same time results support the idea of repeating keywords in the titles one time to get ranked better. Although findings could not confirm the usefulness of long titles, webpages which had repetitive keywords in their titles did not rank well when compared with others. Small changes in titles, such as connecting keywords with "and" can significantly improve ranking.

Experimental results appear to imply that websites that have no outbound links rank better in comparison with others. However, it could not be confirmed that having high quality outbound links can cause websites to rank better. At the same time, results did not find any strong correlation between low quality links and getting ranked more harshly. In addition, not using keywords within anchor text in outbound links and also using indirect outbound links could be helpful.

The experimental studies found that high ranking can be achieved by having a keyword density of around 5% *per search term keyword*. The function of keyword density against ranking is independent of the number of search term keywords.

In summation of the findings, webmasters should avoid having unnecessary outbound links, while attempting to repeat the important keywords of each page one time in their titles to increase the pages ranking in the results page.

6. **References**

[1]   RESTON, VA. comScore Releases January 2011 U.S. Search Engine Rankings. *comscore.* [Online] 2011. [Cited: 8 January 2012.] http://www.comscore.com/Press_Events/Press_Releases/2011/2/comScore_Releases_January_2011_U.S._Search_Engine_Rankings.

[2]   SERVE, IB . Advantages and Disadvantages of Internet Advertising. *article alley.* [Online] 2008. [Cited: 7 January 2012.] http://ibserve.articlealley.com/advantages-and-disadvantages-of-internet-advertising-690918.html.

[3]   Fron, Christine . Internet Advertising Advantages. *Yahoo! Contributor Network.* [Online] 2005. [Cited: 7 January 2012.] http://voices.yahoo.com/internet-advertising-advantages-1497.html.

[4]   Google. Search Engine Optimization starter guide. *Google.* [Online] 2 October 2010. [Cited: 2012 January 7.] http://static.googleusercontent.com/external_content/untrusted_dlcp/www.google.com/en//webmasters/docs/search-engine-optimization-starter-guide.pdf.

[5]   Reston, Va. comScore Releases September 2011 U.S. Search Engine Rankings. *ComScore.* [Online] 2011. [Cited: 7 January 2012.] http://www.comscore.com/Press_Events/Press_Releases/2011/10/comScore_Releases_September_2011_U.S._Search_Engine_Rankings.

[6]   Enge, Eric , et al. *The art of SEO.* [ed.] Mary Treseler. 1st. Sebastopol : O'Reilly, 2010.

[7]   Jerkovic, John . *SEO warrior.* [ed.] Mike Loukides. 1st. Sebastopol : O'Reilly, 2010.

[8]   Falls, Brandon , Goradia, Adi and Perez, Charlene . Google's SEO Report Card. *Google Webmaster Central.* [Online] 2010. [Cited: 7 January 2012.] http://static.googleusercontent.com/external_content/untrusted_dlcp/www.google.com/en//webmasters/docs/google-seo-report-card.pdf.

[9]   Brutlag, Jake . Speed Matters. *Research Blog.* [Online] 2009. [Cited: 7 January 2012.] http://googleresearch.blogspot.com/2009/06/speed-matters.html.

[10]   GoogleWebmasterHelp. Is speed more important than relevance? *Youtube.* [Online] 2010. [Cited: 7 January 2012.] http://www.youtube.com/watch?v=muSIzHurn4U.

[11]   Singhal, Amit and Cutts, Matt . Finding more high-quality sites in search. *google blog.* [Online] 2011. [Cited: 7 January 2012.] http://googleblog.blogspot.com/2011/02/finding-more-high-quality-sites-in.html.

[12]   Singhal, Amit . More guidance on building high-quality sites. *google webmaster central blog.* [Online] 2011. [Cited: 7 January 2012.] http://googlewebmastercentral.blogspot.com/2011/05/more-guidance-on-building-high-quality.html.

[13]   GoogleWebmasterHelp. Does Google consider the URL of an image? *youtube.* [Online] 2009. [Cited: 7 January 2012.] http://www.youtube.com/watch?v=h2SWuUobbr0.

[14]   Lee, Jen and Douvas, Alexi . Ring in the new year with accessible content: Website clinic for non-profits. *google webmaster central blog.* [Online] 2010. [Cited: 7 January 2012.] http://googlewebmastercentral.blogspot.com/2010/12/ring-in-new-year-with-accessible.html.

[15]   Google. BlogHer 2007: Building your audience. *google webmaster central blog.* [Online] 2007. [Cited: 7 January 2012.] http://googlewebmastercentral.blogspot.com/2007/03/blogher-2007-building-your-audience.html.

[16]   —. Image publishing guidelines. *Webmaster Tools Help.* [Online] 2011. [Cited: 7 January 2012.] http://support.google.com/webmasters/bin/answer.py?hl=en&answer=114016.

[17]   Linsley, Peter . Get up-to-date on Image Search. *google webmaster central blog.* [Online] 2009. [Cited: 7 January 2012.] http://.blogspot.com/2009/03/get-up-to-date-on-image-search.html.

[18]   Szymanski, Kaspar , Far, Pierre and Naumann, Sven . Sharing advice from our London site clinic. *google webmaster central blog.* [Online] 2011. [Cited: 7 January 2012.] http://googlewebmastercentral.blogspot.com/2011/04/sharing-advice-from-our-london-site.html.

[19]   Raman. Finding easy-to-read web content. *google blog.* [Online] 2006. [Cited: 7 January 2012.] http://googleblog.blogspot.com/2006/07/finding-easy-to-read-web-content_20.html.

[20]   Google. Link schemes. *Google Webmaster Tools Help.* [Online] 2011. [Cited: 7 January 2012.] http://www.google.com/support/webmasters/bin/answer.py?answer=66356.

[21]   —. Cloaking, sneaky Javascript redirects, and doorway pages. *Google Webmaster Tools Help.* [Online] 2011. [Cited: 8 January 2012.] http://support.google.com/webmasters/bin/answer.py?hl=en&answer=66355.

[22]   —. Hidden text and links. *Google Webmaster Tools Help.* [Online] 2011. [Cited: 8 January 2012.] http://www.google.com/support/webmasters/bin/answer.py?answer=66353.

[23]   —. Automated queries. *Google Webmaster Tools Help.* [Online] 2011. [Cited: 8 January 2012.] http://www.google.com/support/webmasters/bin/answer.py?answer=66357.

[24]   Grappone, Jennifer and Couzin, Gradiva. *Search engine optimization An hour a day.* [ed.] Pete Gaughan. 3rd. Indianapolis : Willey, 2011.

[25]   Michael , Alex and Salter, Ben . *Marketing Through Search Optimization.* 2nd. Oxford : Elsevier Ltd, 2008.

[26]   Bailyn, Evan and Bailyn, Bradley. *Outsmarting Google.* [ed.] Sandra Schroeder, et al. 1st. Indianapolis : Que, 2011.

[27]   Kent, Peter. *Search Engine Optimization For Dummies.* 3rd. Indianapolis : Wiley, 2008.

[28]   Konia, Brad . *Search Engine Optimization with WebPosition Gold.* 2nd. Texas : Wordware, 2002.

[29]   Fox, Vanessa . *Marketing in the Age of Google: Your Online Strategy IS Your Business Strategy.* New Jersey : John wiley & Sons, 2010.

[30]   Ledford, Jerri . *Search Engine Optimization.* [ed.] Mary Beth Wakefield. 2nd. Indianapolis : Wiley, 2009.

[31]   Murray, Glenn. *Seo Secrets.* 2nd. s.l. : Divine Write, 2009.

[34]   Google. Google custom search. *Google.* [Online] 28 April 2009. [Cited: 16 January 2012.] http://www.google.com/cse/.

[35]   Xu, Hui . is Google costume search engine benefits from the same technology that GOOGLE.com has in ranking. *Google custome search.* [Online] 2011. [Cited: 8 January 2012.] https://groups.google.com/a/googleproductforums.com/forum/#!topic/customsearch/764A29s3stg/discussion.