

# Assessing the Legitimacy of a Google Penalty

A Proposal for a Reliable, Robust, and Verifiable Method for Assessing the Legitimacy of a Given Penalty, Demotion, or Ranking Signal

> By Adam Raff and Shivaun Raff Co-founders of <u>Foundem</u> and <u>SearchNeutrality.org</u>

> > January 21, 2013

## **Table of Contents**

Assessing the Legitimacy of a Search Penalty1
Definition of a Penalty1
Factors to be Considered when Assessing a Penalty
Manual versus Algorithmic Penalties3
Google's Most Commonly Cited Rationales
Copied Content3
Lack of Original Content3
A Primary Purpose to Drive Traffic to Other Sites4
Search Engines Shouldn't Feature Other Search Engines4
User Feedback and External Testers4
The Test of a Legitimate Penalty5
The Test of a Legitimate Penalty5 Applying the Test of a Legitimate Penalty to Some Real World Test Cases
Applying the Test of a Legitimate Penalty to Some Real World Test Cases
Applying the Test of a Legitimate Penalty to Some Real World Test Cases
Applying the Test of a Legitimate Penalty to Some Real World Test Cases
Applying the Test of a Legitimate Penalty to Some Real World Test Cases
Applying the Test of a Legitimate Penalty to Some Real World Test Cases
Applying the Test of a Legitimate Penalty to Some Real World Test Cases5Test Case 1: Foundem's June 2006 to December 2010 Search Penalty5Test Case 2: Leading Price Comparison Services Penalised by Panda7Test Case 3: JCPenney9Test Case 4: A Typical Made-for-AdSense (MFA) or Thin-Affiliate Spam Site.10Annex: Background Information about Penalties11

## Assessing the Legitimacy of a Search Penalty

In June 2012, we wrote a proposed framework of remedies<sup>1</sup> designed to end Google's anticompetitive search manipulation practices. We highlighted that the systematic promotion of Google's own services through Universal Search, and the systematic demotion or exclusion of Google's competitors through illegitimately-targeted penalties are two separate mechanisms that will require separate remedies to address. Remedying one of these practices without also remedying the other would not solve the problem; it would simply allow Google to dial-up the un-remedied conduct in order to achieve the same or equivalent anti-competitive effect. If the Commission were to intervene only to prevent Google preferencing its own services, for example, Google could simply re-calibrate its arsenal of penalty algorithms and ranking signals to attain a similar anti-competitive result.

Many of our detailed remedy proposals can be summarised by a single principle:

# Google must be even-handed. It must hold all services, including its own, to exactly the same standards, using exactly the same crawling, indexing, ranking, display, and penalty algorithms.

Adherence to this principle would immediately end Google's ability to systematically favour its own services through the preferential placement and display formats of Universal Search. Adherence to this principle would also end Google's ability to systematically penalise, demote or exclude its competitors. But regulators will require a reliable method for determining when Google is failing to adhere to this principle by embedding insidious anti-competitive factors into its penalty algorithms and ranking signals.

The detection of a penalty is not likely to be a problem; legitimate sites that suddenly find themselves significantly demoted by a Google penalty or algorithm change will usually detect the resultant substantial drop in traffic within hours. The trickier part is how to distinguish legitimate, justifiable penalties and demotions from their illegitimate, unjustifiable, and anti-competitive counterparts.

In our remedy proposals, we suggested that being able to probe and assess Google's claimed *rationale* for a particular change would limit, or even obviate, the need for detailed access to source code. In this document, we expand on this assertion by describing a set of straightforward questions that can be used to provide a robust, reliable, and independently verifiable assessment of the legitimacy of a given penalty, demotion, or ranking signal.

## **Definition of a Penalty**

There probably is no single, all-encompassing definition of a penalty. But, for the purposes of assessing anti-competitive demotions, we can consider a site to be labouring under a penalty when, barring any underlying technical issues, **it consistently ranks significantly lower than its peers for a range of queries for which many users would deem it to be equally or more relevant.** 

For example, the following graphs compare the Google search rankings of several of the U.S.'s leading price comparison services in April 2011, shortly after Google's "Panda" update. Figure 1

<sup>&</sup>lt;sup>1</sup> <u>http://www.foundem.co.uk/Foundem\_Remedy\_Proposals.pdf</u>

shows the small number of un-penalised services, while Figure 2 shows their many recently penalised peers. Note that Google's Price Comparison service is systematically placed at or near the top of Google's search results for all product- and product-price-comparison related queries by Google's Universal Search mechanism, and was therefore entirely unaffected by Panda (as illustrated in Figure 3):



Figure 1: April 2011 Google U.S. search rankings (un-penalised) for NexTag, Pricegrabber, and Shopper across a broad range of product- and product-price-comparison related search terms



Figure 2: April 2011 Google U.S. search rankings (penalised) for Shopping.com, Bizrate, Ciao, Dealcatcher, Dealtime, Kelkoo, Shopmania, Shopzilla, Twenga, and Yahoo! across the same broad range of product- and product-pricecomparison related search terms



Figure 3: April 2011 Google U.S. placement of its own price comparison service (shown in red), Google Product Search, across the same broad range of product- and product-price-comparison related search terms

## Factors to be Considered when Assessing a Penalty

When deciding whether or not a particular penalty or demotion is legitimate, two general factors must always be considered:

- 1. The nature and purpose of the site being penalised, and
- 2. The nature and intent of the queries it is being demoted or excluded for

For example, the criteria for assessing the quality or efficacy of a particular website will vary significantly depending on whether that website is a content publisher (such as a newspaper or advice site), a service provider (such as a vertical search engine or digital mapping service), or an online retailer.

## **Manual versus Algorithmic Penalties**

For a manually imposed penalty, we simply need to consider Google's stated reason and rationale for imposing the penalty. But for automated, algorithmically applied, penalties we need to consider the reasoning and rationale that Google has encoded in its various algorithms and in the selection and weighting of its various ranking signals. Note that this assessment does not require access to the algorithms themselves; it merely requires the ability to ask high-level questions of the Google personnel who are familiar with them.

## **Google's Most Commonly Cited Rationales**

The following are a few examples of general features and characteristics that Google often cites when defending a particular penalty or other change to its algorithms.

## **Copied Content**

"Google says it 'de-indexed' [Foundem] because much of its content – about 87% – was copied from other sites, which it says leads to automatic downgrading in its search results." The Guardian, 20 November 2010<sup>2</sup>

When Google makes such claims, their reasonableness needs to be assessed in the context of the nature and purpose of the site being penalised. Penalising a spam site that copies another site's content wholesale and claims it as its own is justifiable; but penalising a search service that copies content from other sites so that it can efficiently organise, search, and summarise it is not. In short, it is not reasonable to penalise a search service for being precisely what it purports to be: a search service.

Moreover, as Google knows better than most given that it is Google's own business model, a page of search results consisting of individual snippets copied from and attributed to multiple sites with their approval is very different from a page consisting of large swathes of content copied wholesale and unattributed from one other site.

## **Lack of Original Content**

"Google's demotion algorithms...identify sites that systematically suffer from a lack of original content" Google, May 2010

<sup>&</sup>lt;sup>2</sup> <u>http://www.guardian.co.uk/technology/2010/nov/30/google-foundem-ec-competition-rules</u>

*"one change…primarily affects sites that copy others' content and sites with low levels of original content"* Blog Post about Google's Panda update by Matt Cutts, Google's Head of Search Quality, January 28 2011<sup>3</sup>

Whereas "original content" is a reasonable requirement for sites that purport to be content publishers, it is an unreasonable requirement for a search service. Search services are not intended to produce original content, and nor do users expect that they would do so; search services are intended to efficiently organise, search, and summarise the content of others.

## A Primary Purpose to Drive Traffic to Other Sites

"Since the primary purpose of [Foundem] is to drive traffic to other websites, the Quality Team has decided that the initial evaluation was not in error." Google's AdWord Quality Team, August 2006

Once again, context is vital. All search services, horizontal and vertical, exist to deliver users to the other sites that have the content, products or services that the users are searching for:

"Google may be the only company in the world whose stated goal is to have users leave its website as quickly as possible." Google's Mission Statement, prior to September 2009

## Search Engines Shouldn't Feature Other Search Engines

"Google is a search engine. A search engine's job is to point you to destination sites that have the information you are seeking, not to send you to other search engines." Danny Sullivan, The Incredible Stupidity Of Investigating Google For Acting Like A Search Engine, Search Engine Land, November 2010<sup>4</sup>

Such arguments are fallacious, as they overlook the crucial difference between *horizontal* and *vertical* search engines. As Google knows, millions of users every day visit Google specifically looking for vertical search services and the unique information these services provide:

"Vertical search sites are important to [Google] and our users - indeed vertical search sites which offer added value often come top of our search rankings." Google Spokesman, The Register, November 2009<sup>5</sup>

#### **User Feedback and External Testers**

Google often refers to "user polls" or "user feedback" in support of its various actions, without producing any data for independent verification. Amidst the controversy over its recent Panda update, for example, Google referred to its "standard evaluation system" of using "outside testers" to evaluate websites. Google claims that this system is "strictly scientific".<sup>6</sup> But we strongly suspect that it is not.

For Google's evaluation system to be considered scientific, it would need to employ a blind, or preferably double-blind, method, whereby Websites would be stripped of all branding and other

<sup>&</sup>lt;sup>3</sup> <u>http://www.mattcutts.com/blog/algorithm-change-launched/</u>

<sup>&</sup>lt;sup>4</sup> <u>http://searchengineland.com/the-incredible-stupidity-of-investigating-google-for-acting-like-a-search-engine-57268</u>

<sup>&</sup>lt;sup>5</sup> http://www.theregister.co.uk/2009/11/19/google hand of god/

<sup>&</sup>lt;sup>6</sup> http://www.wired.com/business/2011/03/the-panda-that-hates-farms/all/1

identifying features before being assessed. Without this, asking someone who they trust more with their money, American Express or Wazonga.com, for example, cannot be considered a fair, scientific, or legitimate comparison of their respective websites.

## The Test of a Legitimate Penalty

We suggest that the answers to the following questions allow a robust, reliable, and independently verifiable assessment to be made of the legitimacy of a given penalty or demotion when applied to a given site or class of sites for a given query or set of queries. Note that failing any one of these four tests would render the penalty illegitimate.

- 1. Why is this page or site being penalised?
  - a. Is it because of a feature of the page or a feature of the site?
  - b. Is this reason justified within the context of the nature and purpose of the site, and given the nature and intent of the queries the site is being excluded or demoted from?
  - c. Are the site's peers/competitors being held to the same standard?
- 2. What changes to the page or site would need to be made for the penalty to no longer apply?
  - a. Is it reasonable (or even desirable) to require these changes given the nature and purpose of the site, and given the nature and intent of the queries the site is being excluded or demoted from?
- 3. Does Google offer or have a financial interest in a similar service to the one being penalised?
  - a. If so, why is Google's similar service not being penalised in a similar way?
  - b. What features, present or absent, have caused this site to be penalised while exempting Google's similar service?
  - c. Do these differences justify penalising this site and exempting Google's similar service?
- 4. Is the penalty likely to cause frustration or harm to any of Google's users, for example, by preventing them from finding something they are legally and legitimately searching for?

## Applying the Test of a Legitimate Penalty to Some Real World Test Cases

## Test Case 1: Foundem's June 2006 to December 2010 Search Penalty

Shortly after launching its innovative vertical-search service, Foundem was struck by a site-wide algorithmic Google search penalty that systematically demoted its pages beyond the reach of most users. This penalty lasted more than three years, until Google eventually manually intervened to remove the penalty (whitelist Foundem) in December 2010. We have used the various reasons Google has given for this penalty to assess its legitimacy below.

It should be noted that, during its three year penalty, Foundem offered many unique features and provided price comparisons for whole classes of products that were not available elsewhere. For much of Foundem's penalty, for example, Google users searching for price comparisons for motorcycle equipment with queries like "compare prices shoei xr-1000" would not find Foundem's unique price comparisons; instead, they would be presented with dozens of sites that did not offer any kind of price comparison for these products:

hoei XR-1000 Alloy Blu-Sil - car review, road test & used car prices - Mozilla Firefox		🕮 🛲 😻 Shoei XR-1000 - Prices, Photos, Videos - Helmets - Mozilla Firefox	0.0
Edit Yiew Higtory Bookmarks Iools Help		File Edit View History Rookmarks Iools Help	
C X 🟠 🚺 http://www.ciso.co.uk/Shoei_XR_1000_Alloy_Blu_Sil_6512069	습 • 🚺 Google	P C X 🛆 in http://www.wikie.co.uk/guide/shoei-ar-1000-121963.html	😭 🔸 🚮 🖬 Google
Shoei XR-1000 Alloy Blu-Sil - car revi		- 🚫 Shoei XR-1000 - Prices, Photos, Vide	
) Images Video Shopping News Maps More   MSN   Windows-Live	Join Ciao   Login	Help C News Blogs Shopping Videos Top Bl	logs ∺ ∣ Legin
Appliances Books Cameras Cars Computers Electron		Blogs your way 🔘 Shopping 🍚 All News	
Ciao, Ling	Ciao - for people with opinions Earn £1 for your Latest review + Best picks + Write a review + Latest videos First review	Svore Tar Champions League 🗶 • New page 👁	
par internet user,	1	E Shoei XR-1000 Wala > Shoei XR-1000 Ada by Cacitor a Autoray de Acasancies > Hetmels > Shoei XR-1000 Ada by Cacitor	You could write off 75%
nt new online at: 06:50 p.m. on: 17/September/2009 ngratulations! That's why we've just selected you to be the possible winner of an AUDI A3 ( elected, click here: www.selected.winner.co.uk Instant win: FREE FAMILY HOLI	CABRIO JAYI	See Clasconted Herents Free gifts worth 5100 on all Sheel 201000 % from 5199 99 www.bientarase.co.w.	of your debt with an IVA
e - Cara - Motoritika Accessones - Motoritike Halmeta - Shoel XR- Shoel XR-1000 Alloy Blu-Sil Matorika Halmet - Fai face myse Matorika Halmet - Fai face myse Matorika Halmet - Fai face myse	1000 Alloy Bla Sil Cheose action	Cite Nechack     - Glie Nechack     - Studmit a photo     Gregory     Heimes     Specifications: Full Face	£1500 - £3500     €     £3501 - £5000     €     £5001 - £7500     ▼     GO▶
I used this helmet for the couple of mentities and I was quite h though it hadn't visor lists for high speeds, but it was stable tisef while divising my bite. Of course I am not high speed for is it working when speed reaches >160miles Read review	and never opened by wer, so I can'tell how	News New Shoei XR 1000 Symbol helmet	trapped.co.uk Do I qualify?
Rate this product 公公公公公	Pollovine to rattle	Beeding een. 2019 1000 For the upcoming season, Shoel has updated some of the colours on its XR 1000 full face heimet. The	Participate Give feedback Indicate a test
Content of the store of the sto			Submit a photo
		Ads by Gosgle	Send to a friend
		Shoel Motorcycle Helmets Full range of motorcycle helmets Free Delivery & Free spare visor	Top products *
erview Reviews Questions Auctions Resources		www.racevisors.co.uk	1. Airon TR2
iared Links		Shoel Helmets Discourted Shoel helmets. Safety for less!	2 Nolan N43 Air 3 Cabero V2X Carbon
Discounted Helmets Free gifts worth £100 on all Shoei XR1000's from £169.99		www.BkersDirect.co.uk	4. Shark RSF 21
Discounted Heimets Pree gits worth £100 on all Shoel XR1000 \$ from £169.99	www.bikersbargains.co.uk	SHOEI Helmet Sale	5. Shark EvoLine
ine Shoei Visors Tinted, Iridium and black visors Helmets also available www.rac	evisors.co.uk	Shoei XR1000 £199. Great price Shoei X-Spirit £299. Great Service www.HeimetCity.co.ul/ShoeiNeimets	8. Airoh GP
I Helmets Discounted Shoei helmets. Safety for less! www.BikersDirect.co.uk		and a second y of the property of the	2. Nitro N340-VX
shop.com Black, indium, dark, visors Free visor bag offerf www.visorshop.com			Hottest products 🧒
Helmet Sale Shoei XR1000 £199 Great price Shoei X-Spirit £299. Great Senice	www.HelmetCity.co.uk/ShoeHelmets		1. Aral Chaser 2. Shark Evolute
			3. Noter N103
ucts you might be interested in			4. Stori Raid I
tere for under the understand at			5. Arei RX-7 GP

Figure 4: Two Screenshots from September 2009 of two top-ranked sites within Google's search results for the query "compare prices shoei xr-1000". In common with most or all of the sites listed in Google's search results at the time, neither of these sites provides the price comparison the user was looking for. The left hand example features a single eBay shopping listing, a nappy ad, and some sponsored links, while the right hand example features nothing but Google ads.



Figure 5: Screenshot from September 2009 showing Foundem's unique price comparison for the shoei xr-1000 motorcycle helmet. This is presumably what a user typing "compare prices shoei xr-1000" into Google was searching for, yet Foundem's search penalty meant Google's users would never find it.

Question	Answer	Valid?
1. Why is this page or site being	<b>1.</b> Lack of original content and a primary purpose to	No
penalised?	deliver users to other sites	
<b>a.</b> Feature of the page or a feature	a. A feature of the site.	
of the site?	<b>b.</b> No. Foundem is a vertical search service. A lack of	
<b>b.</b> Justified given the nature and	original content and a purpose to deliver users to	
purpose of the site, and the nature	other sites are both defining characteristics of	
and intent of the queries?	search services. Moreover, many of the queries for	
<b>c.</b> Held to same standard as peers?	which Foundem was demoted are explicitly searches	
	for price comparison services. E.g. "compare prices	
	[make model]", "best prices [make model]", and	
	"[make model] prices".	

	<b>c.</b> No. There were many established competitors with similar levels of original content etc. that were not penalised.	
<ul><li>2. What changes would need to be made for the penalty to no longer apply?</li><li>a. Are these changes reasonable?</li></ul>	<ul> <li>2. Foundem would need to change fundamentally from a search engine to a content publisher. Note that none of Foundem's original content, such as user-generated product reviews and ratings or its unique price history graphs, was sufficient to lift the penalty.</li> <li>a. No. It is not reasonable or desirable to insist that vertical search services bloat their pages with original but not necessarily useful information. There are more than enough sites that do this already, e.g. the many thousands of made-for-AdSense (MFA) sites.</li> </ul>	No
<ul> <li>3. Does Google offer a similar service to the one being penalised?</li> <li>a. If so, why is Google's service not being penalised in the same way?</li> <li>b. What features have caused this site to be penalised while exempting Google's?</li> <li>c. Do differences justify different treatment?</li> </ul>	<ul> <li>3. Yes.</li> <li>a. Google does not crawl its own service or subject it to the same ranking or penalty algorithms. Google's services are "blended" into the natural search results via Google's Universal Search mechanism.</li> <li>b. We do not know what the outcome would be if Google switched off Universal Search and instead exposed its own service to the same crawling, indexing, ranking and penalty algorithms as everyone else.</li> <li>c. No.</li> </ul>	No
<b>4.</b> Is the penalty likely to cause frustration or harm to any of Google's users?	<ul> <li>4. Yes. During the course of its three year penalty, Foundem offered features and provided price comparisons for products and whole classes of products that were not available anywhere else. During this time, Google users looking specifically for price comparisons for these products will have wrongly concluded that no such service existed. Users on Yahoo and Bing, where Foundem continued to rank normally, would not have been similarly frustrated.</li> </ul>	No
Conclusion: This penalty fails all four t penalty.	ests comprehensively, and is therefore not a legitimate	

## Test Case 2: Leading Price Comparison Services Penalised by Panda

We have completed the following table based on Google's public statements attempting to explain the seemingly random way in which Panda penalised certain market-leading price comparison services while leaving others (including, of course, Google's own) unaffected.

Question	Answer	Valid?
1. Why is this page or site being	<ol> <li>Lack of original content/surfeit of copied</li> </ol>	No
penalised?	content/borderline spam.	
<b>a.</b> Feature of the page or a feature	<b>a.</b> A feature of the site.	
of the site?	<b>b.</b> No (unless Google can provide additional	
<b>b.</b> Justified given the nature and	explanation). A lack of original content is an inherent	

<ul> <li>purpose of the site, and the nature and intent of the queries?</li> <li>c. Held to same standard as peers?</li> <li>2. What changes would need to be made for the penalty to no longer apply?</li> <li>a. Are these changes reasonable?</li> </ul>	feature of vertical search services. Google has not yet made a convincing argument that explains how these well respected, established sites can be considered borderline spam. <b>c.</b> It seems not. Google has not explained why certain well-regarded price comparison sites have been penalised while others have not. <b>2.</b> To address the lack of original content/copied content, these services would need to change fundamentally from a price comparison service to a content publisher. We do not know what would be required to prevent Google's algorithms from evaluating them as borderline spam. <b>a.</b> No. It is not reasonable or desirable to insist that price comparison services bloat their pages with original but not necessarily useful information, or to penalise them for reproducing snippets of product descriptions from the various retailers they are comparing.	No
<ul> <li>3. Does Google offer a similar service to the one being penalised?</li> <li>a. If so, why is Google's service not being penalised in the same way?</li> <li>b. What features have caused this site to be penalised while exempting Google's?</li> <li>c. Do differences justify different treatment?</li> </ul>	<ul> <li>3. Yes.</li> <li>a. Google does not crawl its own service or subject it to any of the same ranking or penalty algorithms. Google's service is "blended" into the natural search results via Google's Universal Search mechanism.</li> <li>b. We do not know what the outcome would be if Google switched off Universal Search and instead exposed its own service to the same crawling, indexing, ranking and penalty algorithms as everyone else.</li> <li>c. No. Until we know whether Google's price comparison service would be subject to a Panda penalty, and if not, why not, this cannot be considered valid.</li> </ul>	No. (see 3c)
4. Is the penalty likely to cause frustration or harm to any of Google's users?	<ul> <li>4. Yes. Consumer choice has been significantly diminished by this change. This is problematic, as no one price comparison service will excel across all products.</li> <li>Moreover, because Google's own price comparison service now prioritises product listings based on what a merchant is willing to pay Google rather than on the price consumers will have to pay, consumers are inevitably paying higher prices for products than they would if given a natural, broader choice of price comparison services.</li> </ul>	No
services have been substantially dem	a further explanation for why certain leading price comp noted while others have not (and in particular, why its ow ly favoured), the elements of Panda that systematically de cannot be considered legitimate.	'n

#### **Test Case 3: JCPenney**

In a much-publicised example from early 2011, Google applied a manual, three-month search penalty to JCPenney, following allegations that the U.S. retailer had employed prohibited SEO tactics. It was widely reported that, during the penalty, JCPenney no longer appeared anywhere on the first several pages of Google's search results for large numbers of generic queries such as "dresses" or "summer dresses". What was not reported was that this penalty also meant that Google's users could no longer find JCPenney even when using very specific, JCPenney-oriented queries such as "jcpenney dresses" or "dresses at jcpenney".



Figure 6: Two screenshots (from February 14 2011), showing Google search results for two examples of JCPenneyspecific queries – "jcpenney dresses" and "sweater dresses at jcpenney" - during its manually-applied, three-month Google search penalty

If we use our benchmark questions to assess the legitimacy of this penalty we find that it passes tests 1 through 3. But, interestingly, even this seemingly entirely legitimate penalty falls foul of one of the benchmarks (test 4). Due to the nature of Google's site-wide manual penalties, by punishing JCPenney for allegedly attempting to game its algorithms, Google also inadvertently punished significant numbers of its own users—those looking specifically for products at JCPenney. This suggests that Google probably needs to improve the granularity of its manual punitive penalties (if it has not done so already) so that they do not frustrate Google's users. One can also imagine that JCPenney might have grounds to challenge the legitimacy of a penalty that relegates links to its site beyond the reach of most users, even in cases where those users appear to be using Google as a navigational stepping-stone to JCPenney.

Question	Answer	Valid?
<b>1.</b> Why is this page or site being penalised?	1. Alleged cheating	Yes
<b>a.</b> Feature of the page or a feature of the site?	<b>a.</b> Neither – manually applied for alleged	
<b>b.</b> Justified given the nature and purpose of the	cheating	
site, and the nature and intent of the queries?	<b>b.</b> Yes (mainly, see answer 4)	
c. Held to same standard as peers?	<b>c.</b> Presumably.	
2. What changes would need to be made for	<b>2.</b> None. It will expire in 3 months.	Yes
the penalty to no longer apply?	<b>a.</b> n/a	
a. Are these changes reasonable?		
3. Does Google offer a similar service to the	3. Not yet.	Yes
one being penalised?	<b>a.</b> n/a	

a. If so, why is Google's service not being	<b>b.</b> n/a	
penalised in the same way?	<b>c.</b> n/a	
<b>b.</b> What features have caused this site to be		
penalised while exempting Google's?		
<b>c.</b> Do differences justify different treatment?		
<b>4.</b> Is the penalty likely to cause frustration or	4. Yes, when users' search terms	?
harm to any of Google's users?	specifically include "JCPenney", e.g.	
	"JCPenney dresses" or "dresses at	
	JCPenney"	
Conclusion: This is a legitimate penalty, provided the allegations of cheating are true. However, the		

potential to frustrate users searching specifically for products at JCPenney is cause for concern.

## Test Case 4: A Typical Made-for-AdSense (MFA) or Thin-Affiliate Spam Site

Question	Answer	Valid?
<ul> <li>1. Why is this page or site being penalised?</li> <li>a. Feature of the page or a feature of the site?</li> <li>b. Justified given the nature and purpose of the site, and the nature and intent of the queries?</li> <li>c. Held to same standard as peers?</li> </ul>	<ol> <li>Lots of content that has been copied wholesale and unattributed from other sites and lots of low-quality, search-engine-optimised content         <ul> <li>Both</li> <li>Yes. This is a content- rather than service-oriented site. Users don't necessarily know that the articles on this site have either been lifted from somewhere else or have been written more for their ability to rank well in search engines than to impart useful or accurate information.</li> <li>Yes. Its peers are other spam sites.</li> </ul> </li> </ol>	Yes
<ul><li>2. What changes would need to be made for the penalty to no longer apply?</li><li>a. Are these changes reasonable?</li></ul>	<ul> <li>2. The site needs to stop copying content from other sites and presenting it as its own; it also needs to start writing high-quality, accurate content designed to be useful to users</li> <li>a. Yes</li> </ul>	Yes
<ul> <li>3. Does Google offer a similar service to the one being penalised?</li> <li>a. If so, why is Google's service not being penalised in the same way?</li> <li>b. What features have caused this site to be penalised while exempting Google's?</li> <li>c. Do differences justify different treatment?</li> </ul>	3. No. a. n/a b. n/a c. n/a	Yes
<b>4.</b> Is the penalty likely to cause frustration or harm to any of Google's users?	<b>4.</b> No	Yes
<b>Conclusion:</b> This is a legitimate pena	lty.	

## **Annex: Background Information about Penalties**

#### Introduction

A general-purpose, horizontal search engine like Google uses hundreds, or even thousands, of different ranking signals, which it weights and combines to determine the relative placement (rank) of every matching (relevant) search result. In the absence of any political, anti-competitive, or financially-motivated bias, this ordering would reflect the search engine's "best guess" at relevance—that is, the search results that best satisfy a user's query would be listed highest. Exactly how good a particular search engine is at making this guess will depend on the effectiveness of its various interwoven algorithms and the selection, calibration, and relative weighting of its ranking signals.

No two search engines will produce exactly the same search results. In many cases, there is no "right" answer, and different search engines will disagree on the optimum, "best guess" selection and ordering of search results for a given query. But any genuine pursuit of the most relevant results must, by definition, preclude any form of arbitrary discrimination. The problem for Google is that its Universal Search mechanism, which systematically promotes Google's own services, and its increasingly anti-competitive penalty algorithms, which systematically demote or exclude Google's rivals, are both clear examples of financially motivated arbitrary discrimination.

Unfortunately, because of the nature and complexity of search engines and the inherent lack of definitive right and wrong answers, search engine bias is usually difficult or impossible for users to detect; a point eloquently noted by Google's co-founders, shortly before launching their (at-the-time, bias-free) search engine, Google:

"Since it is very difficult even for experts to evaluate search engines, search engine bias is particularly insidious" - Anatomy of a Search Engine, Sergey Brin and Larry Page, April 1988

#### **Penalties and Ranking Signals**

Generally, positive ranking signals serve to improve a page's ranking, while negative signals serve to lower it. A site with a great deal of brand "authority" (an example of a positive ranking signal) like the *New York Times* will tend to rank higher than a site with less brand "authority" such as that of an occasional blogger.

Like most horizontal search engines, Google employs a variety of algorithmic (and manual) penalties—powerful negative ranking signals designed to demote or exclude "undesirable" websites or pages, regardless of how relevant they appear to be to the user's query. But penalties are supposed to be directed at websites that are "undesirable" in a spam sense, or that have been caught "cheating" the search engine's algorithms; they should not be aimed at legitimate services whose only undesirable feature is that Google now views them as competitors to its own growing stable of rival, and often inferior, services.

Search penalties are usually calculated "off-line" (rather than in real time) and are normally applied to an entire site rather than to a particular page. As a result, it is often not possible to work out why a particular page of a website has been penalised; it might have been penalised because of a penalty incurred elsewhere on the site or because of some content or feature of the page that is no longer present.

Because penalties were originally devised as a means of weeding out spam—sites that deliberately disguise themselves as being highly relevant to popular search terms in order to fool search engines into ranking them highly—they tend to be very powerful. Although a penalty signal may be just one signal of thousands that Google uses to determine a page's ranking for a particular search term, the penalty signal is often so heavily weighted that, as far as Google's ranking algorithms are concerned, nothing else matters. For severe penalties, as long as a site remains penalised its pages will never appear anywhere near the top of Google's search results for any query (except, perhaps, for the site's brand-name) no matter how relevant.

Unfortunately, while penalties used to be reserved for spam sites or sites caught attempting to cheat Google's algorithms, Google is increasingly deploying them anti-competitively; targeting algorithmic penalties at characteristics, such as a "lack of original content", that apply to all vertical search services, as well as to certain spam sites.

Google has been overwhelmingly dominant in search and search advertising for over a decade, and there is no sign of this changing anytime soon. As purveyors of the world's overwhelmingly dominant search engine, Google has extraordinary influence over what the World's billions of internet users discover, read, use, and purchase. For the vast majority of online businesses this results in an uncomfortable but unavoidable reliance on Google. It is no exaggeration to say that, with a very few exceptions, any business—large, medium, or small—that is substantially reliant on an online presence can be brought rapidly to the brink of bankruptcy by a Google search penalty. Consequently, an unjustified Google search penalty, whether imposed anti-competitively or in error, has the power to cause grave and irreparable harm to virtually any online business.

Ironically, whereas spammers suffering from old-style spam-targeted penalties can escape the penalty by mending their ways or moving on to a new website, legitimate sites struck by anticompetitive penalties cannot. Short of fundamentally changing their business from a search service to a content publisher, their only means of escape is through manual intervention by Google (e.g. whitelisting).

#### **Manual Intervention**

Traditionally, when Google has wanted to single out a specific site for special treatment, to either manually penalise (blacklist) it or manually immunise (whitelist) it from certain algorithmic penalties, it has done so through the use of manually maintained exception lists, which record the names of specific websites to penalise or exempt.

Distinguishing pure algorithmic promotions and demotions, which act solely on information that Google's algorithms have determined for themselves, from those that act with reference to specific site names or values that have been set by a Google employee, will sometimes be crucial when assessing whether Google is acting in a fair and non-discriminatory manner.

"Panda" and its various follow-on updates may have blurred the line between what can and cannot be considered a "manual" intervention. Whether this blurring was intentional or not, we probably need to extend our definition of a manual intervention. For example, manually devising, adjusting, and calibrating a number of different ranking signals until an algorithm produces a desired outcome, such as automatically demoting or promoting a particular website or group of websites, probably needs to be considered a form of manual intervention.