

Rating Guidelines	3
1.0 Welcome to the Ads Quality Evaluation Program!	3
1.1 Ad Rating Overview	3
1.2 Important Rating Definitions and Ideas	3
1.3 The Purpose of Ads Quality Rating	5
1.3.1 Google Search Engine Screen Shot	5
1.3.2 Google Search Results Screen Shot	6
1.4 Raters Must Understand the User	6
1.5 Internet Safety Information	6
1.6 Confidentiality While Performing Your Duties	7
1.6.1 Entering Email Addresses	7
1.6.2 Communicating with Advertisers	7
1.6.3 Identifying Yourself as a Rater	7
1.7 Browser and System Software	8
1.8 Guidelines for Different Rating Countries	8
2.0 Researching the Query	9
2.1 Classification of User Intent: Action, Information, and Navigation – “Do-Know-Go”	9
2.1.1 Action Queries – “Do”	10
2.1.2 Information Queries – “Know”	10
2.1.3 Navigation Queries – “Go”	10
2.2 Understanding the Query and User Intent	10
2.2.1 Rate According to the User Intent, not the Query Words	11
2.2.2 Queries with Multiple Meanings	12
2.2.3 Spelling Mistakes	13
2.2.4 Unclear Intent	13
2.2.5 Incomprehensible Queries	14
2.2.6 Timeliness	14
3.0 Using EWOQ: Introduction	15
3.1 Accessing the EWOQ Ads Evaluation System	15
3.2 Rating a Task	15
3.3 Acquiring Tasks: the Ads Rating Home Page	15
3.3.1 Ad Rating Home Screen Shot	15
3.4 Starting to Rate: the Ad Rating Task Page	17
3.4.1 Ad Rating Task Screen Shot	17
3.5 Reviewing and Re-rating Unresolved Tasks: Task Deliberation	20
3.5.1 Ad Rating Task in Deliberation Screen Shot	20
3.5.2 The Purpose of Task Deliberation	22
3.5.3 Administrator and Moderator Comments	24
3.6 Managing Open Tasks	24
3.6.1 Task Management: Ad Rating Home Screen Shot	24
3.7 Prioritizing your Tasks	25
3.7.1 Tools to Help Manage Your Work	26
3.8 Task Deliberation Etiquette	27
4.0 Rating a Task	28
4.0.1 Using the Slider Bar	28
4.1 Ad Creative	29
4.1.1 Ad Creative Overview – What You’re Doing	29
4.1.2 Ad Creative Flag Overview	30
4.1.3 Ad Creative Flag Reference	30
4.1.4 Ad Creative Rating Reference Chart	33-34
4.2 Landing Page	35
4.2.1 Landing Page Overview – What You’re Doing	35
4.2.2 Landing Page Flag Overview	36
4.2.3 Landing Page Flag Reference	36
4.2.4 Landing Page Reference Chart	41-42
4.3 Overall Quality	43
4.3.1 Overall Quality Overview – What You’re Doing	43
4.3.2 Overall Quality Scale Ratings	44
4.3.3 Overall Quality Rating Reference Chart	46
4.4 Gut Reaction	47
4.4.1 Gut Reaction Overview	47
4.4.2 Gut Reaction Reference	48
5.0 Comment Codes	49

5.1	Evaluation Comment Codes	49
5.1.1	Advertiser Interpretation based on a Secondary Interpretation	49
5.1.2	Advertiser Interpretation based on Spelling Corrections	50
5.1.3	Landing Page Codes	50
5.1.4	Overall Quality Codes	51
5.1.5	Communication Comment Codes	52
5.2	Using Comment Codes	53
6.0	Common Situations and Special Cases	54
6.1	Pornography, Escort Services, and Sexual Services	54
6.2	Aribtrage, Made For Ads Sites, Etc	55
6.3	Fake Music/Video/Software Sharing Sites	56
6.4	Malware Sites	59
6.5	Information Harvesters	59
6.6	Too-Good-To-Be-True Offers and Sucker Deals	59
6.7	Fake Review Sites	60
6.8	When You Think Something is Illegal	60
7.0	Experimental Projects	62
7.1	Experimental Projects Introduction	62
8.0	Addendum: Differences in Current Guidelines Versus Older Guidelines	63

General Guidelines for Ads Quality Evaluation

1.0 Welcome to the Ads Quality Evaluation Program!

As an Ads Quality Evaluator, you will study what search engine users are looking for, examine the advertisements we show alongside search results for these queries, and evaluate the usefulness of these ads for the search engine user. Your evaluations directly contribute to maintaining the high quality of ads presented on Google search pages.

Please take the time to carefully read through these guidelines. Once you have become very comfortable with the concepts presented, you will be ready to begin evaluating ads.

1.1 Ad Rating Overview

As an Ads Quality Evaluator (AQE) you examine queries entered by users and analyze the advertisers that are paired with them. As an AQE, your unit of work is a task. Your analysis contains several parts. For each task, you will:

- Conduct research in order to understand the query and what the user was looking for (User Intent)
- Evaluate the usefulness and relevance of the advertisement for the user (the Ad Creative rating)
- Evaluate the page the ad leads to if the user were to click on it (Landing Page rating)
- Evaluate and rate the quality of the advertiser's site overall (Overall Quality rating)
- Choose a personal, subjective rating of the advertising experience (Gut Reaction rating)
- Determine if the advertiser fits into one or more other defined categories

There are other things to consider as you perform your ad quality rating duties, but if you focus on these things, and do them well, you will do a great job.

1.2 Important Rating Definitions and Ideas

Search Engine: A search engine is a website that allows users to search the Web by entering words or symbols into a search box. The most popular and trustworthy search engines reliably provide objective, useful search results that help users find what they're looking for.

Query: A query is the set of words, numbers, and/or symbols that a user types in the search box of a search engine. We will sometimes refer to this set of words, numbers, or symbols as the "query terms". Some people also call these "key words". In these guidelines, queries will have square brackets around them. If a user types the words digital cameras in the search box, we will display: [digital cameras].

Ad Creative: The ad creative is an advertisement that appears under the "Ads" heading on the Google search results page. Traditional ad creatives have a title, a body, and a visible URL (see example below). New ad creatives have additional elements like maps, videos, extra links, etc. When a user clicks on an ad creative, it takes him or her to the advertiser's landing page. We also refer to ad creatives as creatives or ads. A traditional ad creative looks like this:

The Official Honda Site	↔Title
See new Honda pictures & specs.	↔Body
Get a Free Dealer Quote today.	↔Body
www.honda.com	↔Visible URL

Landing Page: The landing page is where the user is taken when he or she clicks on the ad creative. This is the advertiser page that you evaluate for your landing page rating. Landing pages can be homepages (i.e. www.apple.com) or other pages within an advertiser's site that is focused on what the user is looking for (i.e. www.apple.com/iphone). In ads quality evaluation, the ad creative and landing page are always evaluated independently of each other.

Landing Site: The landing site is the entire site and all subpages accessible (via links, buttons, search boxes, etc.) through the landing page. Links to other sites with different URLs are not part of the same landing site.

Advertiser: For the purposes of ads quality evaluation, an advertiser is a company or individual who chooses to show advertisements to users of the Google search engine.

Task: This is the basic unit of work in ads quality rating. A task has two main components: 1) the user query; and 2) the advertiser.

User Intent: A user's intent is what the user is really seeking. Some queries perfectly represent what the user intent is while others are more difficult to decipher. Your first responsibility when evaluating a task is to do your best to determine the user intent behind the query.

Search Results / Natural Search Results: Google and other major search engines respond to user queries by providing relevant, objective search results. Because advertisers can't buy better placement or otherwise modify what a search engine provides, the results of the search are called natural or organic search results. Most search engines display advertisements on their search results page, but they are careful to label them ads so users can differentiate between natural search results and paid advertisements.

Ads Quality Rater: Ads quality raters use a web-based system to acquire and evaluate ad rating tasks. We use the terms ads quality rater, rater, and ads quality evaluator (AQE) interchangeably.

Ads: For the purposes of ads quality evaluation, the advertisements that appear along the right-hand side of Google's search results page (and occasionally at the top of the page). They are explicitly labeled as ads and appear when an advertiser thinks their ad will be of interest to a user who entered a query into the Google search engine.

URL: The URL is a unique Web address, such as <http://www.ford.com>. In ads quality evaluation, you will see a URL at the bottom of the ad creative. Each landing page has a unique URL. Occasionally, you may encounter a user query that is a URL, i.e. [www.ford.com].

Visible URL: The visible URL is the web address shown at the bottom of the ad creative. The landing page URL is not always the same as the visible URL: sometimes advertisers display their homepage in their ad creative but take users who click on the ad to specific pages within their site. For example, an ad for McDonald's might display <http://www.mcdonalds.com> as their visible URL but actually take the user to a specific page like this one: http://www.mcdonalds.com/usa/rest_locator.html.

Rating Language and Rating Country: All tasks have a rating language and rating country associated with them. You will be able to choose the rating language and country you were hired to work with.

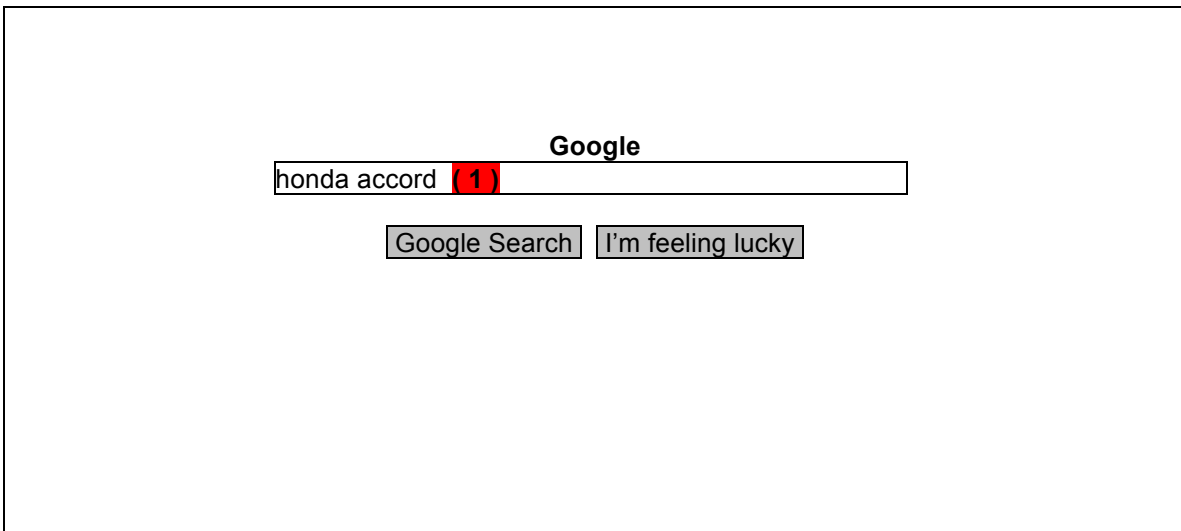
Scale Rating: The scale ratings and how to use them will be described in detail in Section 4: Rating a Task, but here is a brief overview. There are four scale ratings: two ratings are based on the user's query and how the advertiser responds to it; one rating is based on your analysis of the advertiser's site overall and without regard to the query; a final rating represents the overall advertising experience from your personal perspective.

1.3 The Purpose of Ads Quality Rating

Your ratings will be used to evaluate the quality of search ads (the ads that appear as ads on Google's search results page) and how useful these ads are for Google users. The following two screen shots illustrate what you evaluate in this project.

The user enters a query into the Google search engine (1.3.1). The Google search engine returns results (1.3.2) based on the search terms, or query **(1)**, entered by the user. Natural search results **(2)** are displayed along the left-hand side. Advertisements **(3)** are listed on the right-hand side under the "ads" heading. Ads quality evaluation measures the relevance and quality of the ads shown here, and your efforts help us improve the quality of ads displayed to our users.

1.3.1 Google Search Engine Screen Shot



1.3.2 Google Search Results Screen Shot

Google

Web results Results 1-10 of ... for honda accord

<p>2009 Honda Accord Sedan - Official Honda Web Site (2) See reviews, features and specs of the 2012 Honda Accord Sedan at the official Honda Web site. View pictures and learn about the latest changes for 2012. automobiles.honda.com/accord-sedan/ - 53k - Cached - Similar pages</p> <p>Honda Accord Family - 2012 - Official Honda Web Site The Accord combines traditional Honda standards of quality, durability, reliability, environmental friendliness and value. Choose the sporty Coupe or the ... automobiles.honda.com/accord/ - 23k - Cached - Similar pages</p> <p>Honda Accord - Wikipedia, the free encyclopedia The Honda Accord is an automobile manufactured by Honda since 1976, debuting as a compact hatchback and evolving into a intermediate vehicle. ... en.wikipedia.org/wiki/Honda_Accord - 186k - Cached - Similar pages</p> <p>...</p>	<p>Ads (3)</p> <p>2012 10 Top Cars 244 HP, EPA 20 mpg / 29 mpg New V6 Trim Level, Top Quality www.Edmunds.com</p> <p>Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt www.ut.labore.et</p> <p>Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt www.ut.labore.et</p>
---	--

1.4 Raters Must Understand the User

It is very important for you to understand the user. The user is someone who lives in your rating country and reads your rating language and has typed the query into the search box. As a rater, you will imagine the user is looking at a page of search results and one or more ads.

Google wants to offer high-quality ads that are relevant to what users are looking for. Your analysis helps us understand both the quality of ads and their relevance to what users are looking for.

You must be very familiar with the rating country and rating language in order to represent the experience of users in your rating country. If you do not have the knowledge to do this, please inform your employer.

1.5 Internet Safety Information

In the course of your work, you will visit and evaluate many different web pages. While it is rare to encounter an advertiser landing page that may harm your computer, it is possible. Your computer may come to harm unless you are careful. Never download any executables, applications, or other potentially dangerous files, or click on any links that you are uncomfortable with. **We**

strongly recommend that you have antivirus and anti-spyware protection on your computer. This software must be updated frequently or your computer will not be protected. There are many free and for-purchase antivirus and anti-spyware products available on the Internet.

Here are links to Wikipedia articles with information about antivirus software and spyware:

http://en.wikipedia.org/wiki/Antivirus_software

<http://en.wikipedia.org/wiki/Spyware>

Some advertisers may offer to download applications or other files. Such downloads are not necessary to fully evaluate the task since your evaluation primarily involves viewing the advertiser's landing page and site. It **never** involves downloading software from the advertiser.

Note: it is generally safe to accept a "security-alert / invalid-certificate" on a limited session basis when prompted. This type of alert is basically just a warning that the encryption may not be secure. Clicking the following links will show what these kinds of alerts look like:

http://help.its.maine.edu/images/help/computing/casecurityalertie/image_preview

<http://help.its.maine.edu/images/help/computing/casecurityalertfirefox/image>

1.6 Confidentiality While Performing Your Duties

Remember that **the work you are doing is confidential**. There are a few things you should keep in mind as you perform your duties.

1.6.1 Entering Email Addresses

Never enter real personal information on any page you encounter in the course of an Ads Quality Evaluation project, even if it is clearly legitimate. You may sometimes be tempted to enter a fake address to gain access to a site if you are not sure whether it is legitimate or not. If you have a throwaway email address that cannot be traced to you personally, feel free to use it for that purpose. However, you should never enter a made-up email address that has a real domain after the @ symbol: for example, fakeaddress@hotmail.com. There is no way for you to be sure you're not using a real address of someone else. Don't do it.

1.6.2 Communicating with Advertisers

You may sometimes be tempted to contact advertisers to learn information that will help you decide on a rating. **Don't do this under any circumstances**, whether by phone, email, chat, comment box on the landing site, or any other means.

1.6.3 Identifying Yourself as a Rater

Never, under any circumstances, **give any information to an advertiser that even hints that you are an ads evaluator**, or that provides information of any sort about Ads Quality Evaluation. Do not make any allusions to ads evaluation terminology, policies, or personages in "comment here" or "user feedback" boxes of an advertiser's website. (You shouldn't be using these at all!)

1.7 Browser and System Software

There are no particular requirements for what kind of operating system or web browser you should use. All that is necessary is that you use a browser that is capable of properly displaying the EWOQ interface (and most browsers are).

While it is not very common, some landing pages display different content to different browsers. You may encounter situations where raters who use one browser see one thing, and users who use a different browser see another. In cases like this, all raters should base their ratings on what shows up on Internet Explorer for Windows. Raters who are unable to view the site under those conditions should reject the task.

This does not happen very often, so you should not hesitate to use Chrome, Firefox, Opera, Safari, or another browser if you prefer.

1.8 Guidelines for Different Rating Countries

While ads quality evaluation is conducted in several languages, **the guidelines are the same for all languages and countries**. Currently, the ads quality evaluation guidelines are available only in English. Many of the query and ad examples used in the guidelines are also queries and ads from the United States of America. Rating examples from other languages can be found on the Rater Hub. The guidelines should be applied consistently regardless of your rating country or language. You will have the opportunity to ask language or country specific questions when you begin evaluating tasks in your rating language.

2.0 Researching the Query

Your primary objective as a researcher is to review user queries and decide if the ad we pick is relevant and useful for that user. Before evaluating any task, you must understand the user query. Some user queries will be very easy for you to interpret and will require no additional research. Others may require research on your part: you may need to do a web search, or look up an unfamiliar term in a dictionary. We encourage you to use web search engines, online maps, dictionaries, or encyclopedias to help you understand the user intent.

Here are some tools raters may find useful for conducting query research:

- www.google.com – Google search engine
- <http://www.google.com/support/websearch/bin/answer.py?answer=136861> – advanced Google search techniques
- babelfish.yahoo.com and translate.google.com – language translation
- www.wikipedia.org – online encyclopedia

You are not limited to these tools. Feel free to use whatever resource works best for you.

You and other raters may disagree about how to interpret some user queries. Agreement between raters is not necessary in order to evaluate a task. Try to be reasonable with your query interpretation: implausible or outlandish query interpretations aren't going to help Google provide better ads to its users.

If the query is meaningless even after doing research, it means that there is probably no relevant advertisement that can be paired with the query. This situation is explained in Section 2.2.5, below.

2.1 Classification of User Intent: Action, Information, and Navigation – “Do-Know-Go”

Sometimes it is helpful to classify the user intent for a query in one of these three categories:

Action intent - The user is trying to accomplish a goal or engage in an activity, such as download software, play a game online, send flowers, find entertaining videos, etc. These are “do” queries: the user wants to do something.

Information intent - The user is trying to find information. These are “know” queries: the user wants to know something.

Navigation intent - The user is trying to navigate to a website or webpage. These are “go” queries: the user wants to go to a webpage or website.

An easy way to remember this is “Do-Know-Go”. Classifying queries this way can help you figure out what a user wants to do and what—if any—advertiser would be an acceptable match for them. Please note that many queries fit into more than one type of user intent. Below are some examples of queries of each type for a user from the United States.

2.1.1 Action Queries - “Do”

The intent of an **action query** is to accomplish a goal or engage in an activity on the Internet. The goal or activity is to download, to buy, to obtain, to be entertained by, or to interact with a resource that is available on the Internet.

Query	User Intent	URL of a Useful Page	Description of a Useful Page
[geography quiz]	Take an online geography quiz	http://www.lufthansa-usa.com/useugame2007/html/play.html	Page with an online geography quiz that the user can take
[Citizen Kane DVD]	Purchase this DVD	http://www.amazon.com/Citizen-Kane-Georgia-Backus/dp/B00003CX9E http://www.cduniverse.com/productinfo.asp?pid=1980921	Pages on which to purchase this DVD
[buy flowers]	Order flowers online	http://www.ftd.com http://www.1800flowers.com http://www.proflowers.com	Pages on which to order flowers online

2.1.2 Information Queries - "Know"

An **information query** seeks information on a topic. The user wants to **know** something; the goal is to find information.

Query	User Intent	URL of a Useful Page	Description of a Useful Page
[how to remove candle wax from carpet]	Find information on how to remove candle wax from carpet	http://homeparents.about.com/cs/householdtips/ht/Candle_Wax.htm	About.com page about dealing with candle wax.
[cryptology use in WWII]	Find information about how cryptology was used in World War II	http://www.nationalmuseum.af.mil/factsheets/factsheet.asp?id=9722	U.S. Air Force Museum article about cryptology use during WWII
[Switzerland]	Find travel and tourism info for planning a vacation, or find information about Swiss economy, geography, or languages	http://www.lonelyplanet.com/worldguide/destinations/europe/switzerland https://www.cia.gov/cia/publications/factbook/geos/sz.html	Travel guide on Switzerland Informative CIA webpage on Switzerland

2.1.3 Navigational Queries - "Go"

The intent of a **navigation query** is to locate a specific webpage. The user has a single webpage or website in mind. This single webpage is called the **target** of the query. The user wants to **go** to the target page. Here are some examples:

Query	User Intent	URL of a Useful Page	Description of a Useful Page
[ibm]	Navigate to the IBM homepage	http://www.ibm.com	Official homepage of the IBM Corporation
[youtube]	Navigate to the YouTube homepage	http://www.youtube.com	Official homepage of YouTube
[yahoo mail]	Navigate to the Yahoo! Mail login page	http://mail.yahoo.com	Yahoo! Mail login page

2.2 Understanding the Query and User Intent

You will base your ad creative and landing page ratings on how well the creative and the landing page match the **user intent behind the query**. What do we mean by "user intent", and how do

you figure out what it is?

A user enters a query into Google because he or she wants something. Maybe the user wants information; for example, the name of the president of South Korea, or information about heart disease. Maybe the user wants to do something: for example, purchase a camera, or play an online game. Maybe the user wants to go to a specific website or web page: for example, the YouTube homepage, or Yahoo! Mail.

There are all kinds of things a user might want. But always assume that the user wants **something**. The **user intent** is whatever the user wanted. Always do your best to put yourself in the user's place and figure out what would satisfy his or her intent.

This means that it is **not good enough** to look at the words of the query, and check whether they are repeated in the ad or the landing page. You need to figure out, to the best of your ability, what the user wants, and decide whether the creative and landing page will satisfy that.

Don't worry if it's not always clear which category a query belongs to. For example, take the query

[youtube pachelbel canon funtwo]

It's very clear what the user wants. (At least, it's clear after you do some research! This might be a very puzzling query if you didn't research it.) The user wants to look at a certain YouTube video of a young guitarist playing Pachelbel's Canon. Now, is this a Navigational query (since there is a specific page that the user wants to go to) or an Action, or "do", query (since the user wants to watch a particular video)?

The answer: it is not important! If you know exactly what the user wants, then you are in good shape to decide what he or she will think of particular ad creatives or landing pages. These query categories aren't perfect, and we only provide them to give you a way to start thinking about what the user intent really is.

2.2.1 Rate According to the User Intent, not the Query Words

The most common mistake is also the worst mistake you can make: basing your rating just on the words of the query, without bothering to understand the user intent.

Here is an example:

[training table provo utah]

Here's an ad creative you might see:

Provo Table Warehouse
Buy all kinds of tables. Provo's oldest
table and chair retailer!
www.provo-table-warehouse.com

If you don't bother to research the query, you might think this sounds like a fantastic ad. It sounds like the user wants to buy something called a "training table" in Provo (a city in the state of Utah in the United States).

Try a Google search on this query, though. It turns out that this is the name of a restaurant chain. The user is almost certainly interested in the restaurant. This ad for tables is almost certainly no good at all.

Here's another example, one we used above:

[youtube pachelbel canon funtwo]

Let's say we see the following ads:

Pachelbel's Canon

Buy classical music recordings
on Amazon!
www.amazon.com

Canon Printers

Best prices, full warranties
www.printer-paradise-warehouse.com

Funtwo

Get all the best **funtwo** deals
funtwo sales and coupons
www.ebay.com

Again, this is very clearly a search for a particular, popular YouTube video of a guitarist who goes by the name "funtwo". None of these ads have much to do with that user intent. The first one is at least marginally related, but really doesn't have much to do with the user intent. The second one repeats a word from the query, but has absolutely nothing to do with the user intent. The third also repeats a word from the query, but makes pretty much no sense at all.

None of these are good ads for the user; but if you're not bothering to research the query and figure out the user intent, you might imagine that they were fine, just because they repeated a few words from the query. This is a very bad mistake. If you don't understand the user intent you can't determine how well the advertiser responds to it. We encourage you to take query research seriously, and be prepared to discuss your research with others.

2.2.2 Queries with Multiple Meanings

Many queries have more than one meaning. For example, the query [apple], in the United States might refer to the computer brand, the fruit, or the music company. We will call these possible meanings query interpretations.

Dominant Interpretation: Sometimes, there is one query interpretation that most users would agree to. We will call this meaning the **dominant interpretation**. For example, most users typing [windows], in the USA want results on the Microsoft operating system, rather than the glass windows found on walls. The dominant interpretation should be clear to you, especially after doing a little web research.

Common Interpretations: In some cases, there is no dominant interpretation. The query [mercury], in the USA might refer to the car brand, the planet, or the chemical element (Hg). While none of these is clearly dominant, all are **common interpretations**. People might want to see search results related to any of these interpretations.

Minor Interpretations: Sometimes you will find less common interpretations. We will call these **minor interpretations**. Consider again the query [mercury]. Possible meanings exist that even most users in the USA probably don't know about, such as Mercury Marine Insurance and the San Jose Mercury News. These are minor interpretations of the query, [mercury].

We call a query **ambiguous** when there are several plausible common query interpretations.

In general, keep **all** the plausible meanings in mind. In every task you do, you will be looking at a particular ad and landing page. If that ad and landing page are relevant to **one** of the possible meanings, go ahead and assume that that's the meaning the user intended.

Keep in mind only the **dominant** or **common** interpretations unless there is a compelling reason to consider a **minor** interpretation. You should use your common sense to decide what meanings are more plausible and less plausible. Almost any word can mean more than one thing if you really stretch, but that's not very helpful when you're trying to judge what a real person actually wanted when they typed in that query. For example, if someone enters the query [paris], they almost certainly mean Paris, France. Don't decide that the user might have meant Paris, Texas, or Paris Hilton. If the user intended one of these other interpretations, he or she probably would have entered some additional text to specify it. This can be a judgment call, and you and other raters might have legitimate differences of opinion of what a user meant. In general, try to use your common sense and put yourself in the place of the user— don't spend time on outlandish or perverse interpretations of the query.

Most of the time advertisers will respond to dominant and common interpretations of a query. Occasionally, however, advertisers will respond to minor interpretations of the query. When this happens, you must lower your scale ratings to account for this. In Sections 5.1.1 and 5.1.2, we discuss in detail how to rate pages when the advertiser responds to minor interpretations.

When an advertiser interprets a query in a way that doesn't make sense or is likely to be meaningless for the user, there is no reason for you to consider it a legitimate interpretation at all. It is very important for you to identify and correctly evaluate implausible, meaningless, and useless query interpretations: such interpretations require specific negative ratings.

2.2.3 Spelling Mistakes

Users often misspell words and make typos. In general, this **does not change how you interpret the query**. If you can understand what the user meant to type, you can still figure out the user intent. Remember, it's Google's job to give people the search results they're looking for and show them ads that will interest them, and it's your job to figure out if Google is doing that. It might be harder for Google to do this when the user misspells or mistypes a query, but that doesn't matter: whether it's hard or easy, Google should provide relevant, useful ads, and you need to tell us whether we're doing this or not.

It is very important to research the query as the user entered it. You should not assume the query is misspelled because the search engine provides a spelling correction suggestion, or because the advertiser assumes a different spelling is correct.

When an advertiser assumes the user made a spelling mistake and serves an ad based on the advertiser's spelling correction, you may need to note this when you submit the task and adjust your scale ratings. Section 5.1.2 tells you what to do in this situation.

2.2.4 Unclear Intent

Sometimes it's very clear what the query is about, but not clear exactly what the user wants to do. Say the user enters this query:

[canon powershot tx1]

One possibility is that this is an **action** “do” query, and the user intent is to buy this particular camera. Another possibility is that this is an **informational** “know” query, and the user wants to read reviews or get information about the camera. How should you proceed?

If both (or all!) the possibilities seem completely plausible and normal, **keep them all in mind**. An ad or landing page can deserve good ratings if it satisfies **any** of the reasonable possible user intents. For this query, an ad for a site where you can buy this camera and an ad for a site that gives a review of this camera are **both** good.

A query like

[spain]

is very similar. Maybe the user wants to learn about the history or culture of Spain; maybe the user wants to vacation there. There is no way to tell just from the word "Spain". Either kind of ad might be good.

As with queries with multiple meanings, use your common sense. Don't worry about coming up with every possible intent that could have been behind a query, and don't give good ratings to sites that are only good given an outlandish or unlikely interpretation of the user intent. Remember, your goal is to give us your best guess about what the user actually thought--not to find some reason to give a good (or bad) rating. In the case of "Spain", for example, don't give good ratings to ads for recordings of "Lady of Spain" or people-search sites offering information about people whose last name is Spain.

2.2.5 Incomprehensible Queries

If you simply have no idea what the query is about, even after fully researching it, then post a comment explaining this and **give the lowest scale ratings for ad creative and landing page**. The only thing that makes an ad or landing page deserve good ad creative and landing page ratings is a clear connection to the user intent. Don't assume that just because an advertiser is showing an ad that the advertiser has correctly understood the query. If there is no reason to think that the ad or landing page are related to the user intent, **assume it's a bad ad**. If a trained quality evaluator can't determine if it's a good ad, then it probably isn't.

2.2.6 Timeliness

A query may be interpreted differently at different points in time. In 1994, a user in the United States who typed [President Bush] was looking for information on President George H.W. Bush. In 2011, his son, former President George W. Bush, is the more likely interpretation. You should always rate according to the current interpretation, unless the query clearly specifies otherwise.

3.0 Using EWOQ: An Introduction

Welcome to EWOQ, the ads evaluation system you will use as an ads quality evaluator. You will acquire tasks and rate them based on the guidelines given to you. For regular ads evaluation, a **task** has two components: a **query** and an **advertiser**. As you work in the EWOQ interface, you will acquire tasks as you need them and submit your ratings as you complete them.

3.1 Accessing the EWOQ Ads Evaluation System

To access the EWOQ system, you must first log into your Google account. The following link takes you to EWOQ:

<https://www.google.com/evaluation/ads/search/rating/home>

You must supply your Google Account Id and password if you are not already logged in.

3.2 Rating a Task

In general, rating a task involves the following steps:

- Acquiring Tasks: the Ad Rating Home Page (See Section 3.3)
- Starting to Rate: the Ad Rating Task Page (See Section 3.4)
- Reviewing and Re-rating Unresolved Tasks: Task Deliberation (See Section 3.5)
- Managing Open Tasks (See Section 3.6)
- Acquiring More Tasks

3.3 Acquiring Tasks: the Ads Rating Home Page

3.3.1 Ad Rating Home Screen Shot

The screenshot shows the EWOQ Ads Rating Home Page. At the top left, the text "ad rating" is followed by a red box containing the number 1. At the top right, the text "jane.doe [ad rating – experimental – logout]" is displayed, with red boxes containing the numbers 2, 3, and 4 pointing to "ad rating", "experimental", and "logout" respectively. Below this, the text "Ads Rating Home" is followed by a red box containing the number 5. A navigation bar contains several links: "acquire next ad rating task" (6), "Any Language" with a dropdown arrow (7), "history" (8), "guidelines" (9), and "hub" (10). Below the navigation bar, there are five sections, each with a heading and a red box containing a number: "Acquired Tasks" (11), "Unrated or Partially Rated Tasks" (12), "Held Open by Admin" (13), "Resolving" (14), and "Resolved but not Closed" (15). Each section is followed by the text "No tasks in this category...".

The red numbers represent the following:

1. **ad rating**
This text shows what type of Task you are working on. In this case, the Task type is “ad rating”.
2. **jane.doe@gmail.com**
Your GMAIL account.
3. **ad rating – experimental**
Clicking on “ad rating” takes you to the Ad Rating Home page. Clicking on “experimental” takes you to the Experimental Projects Home page.
4. **logout**
Click on this link to end your EWOQ session. Please LOGOUT when you are done with your work session.
5. **Ads Rating Home**
Shows you where you are inside the EWOQ system. This is where you acquire Tasks and where you can see your Task list.
6. **acquire next ad rating Task**
Click this button to acquire Tasks.
7. **language drop-down menu**
You can select a language from this drop-down menu. The default setting, “any language”, gives you a task from any language to which you are assigned. If you want to acquire tasks from only one language, choose the language here.
8. **history**
This report shows your ratings history for a chosen period.
9. **guidelines**
Click on this link to read the “Ads Quality Evaluation Guidelines”.
10. **hub**
Link to the Rater Hub, a resource for ads quality raters.
11. **Acquired Tasks**
Shows whether Tasks have been acquired. Tasks are in one of four different categories, depending on what actions have been taken with a task. Tasks are listed by row, with task information in each of several columns (see Section 3.7.1 for ways to use this information effectively). When there are no Tasks in a category, a note will say “No Tasks in this category...”
12. **Unrated or Partially Rated**
Acquired tasks in this category have not been rated, or you started rating the task but saved a draft rather than submitting it. When you submit a task it disappears from this category. When you finish your workday, you should submit or release all unrated tasks.
13. **Held Open by Admin**
Acquired tasks in this category have been rated and submitted, but an administrator has determined that discussion or a rating adjustment from you or other team-members on the task is required. Administrators can hold open Resolving or Resolved but not Closed tasks. Tasks in this category should be reviewed as soon as possible.
14. **Resolving**
Acquired tasks in this category have been rated and submitted by you and other team members on the task. The EWOQ system analyzed the ratings of everyone who evaluated the task and determined there is a substantial difference of opinion about the ratings or flags for the task. Further discussion is needed. Tasks in this category should be reviewed at least once a day. You are required to discuss all Resolving tasks.
15. **Resolved but not Closed**
Tasks in this category are moved from the Resolving or Held Open by Admin categories when the Task’s issues in dispute have been resolved. You may view and change ratings or flags, or post comments on these tasks until they are closed by administrators.

3.4 Starting to Rate: the Ad Rating Task Page

3.4.1 Ad Rating Task Screen Shot

ad rating → ad rating task jane.doe [ad rating – experimental – logout]




(1)


Ad Rating Task – football games (2)

release task (3) - reject task (4)

Query	football games (5) research this page (optional)
Language	English-ads (6)
Query Description	None (7)
Landing Page	visit landing page (8)
Due Date	None (7)
Ad Creative	Free Sports Games (9) Play Free Sports Games at Games.com Choose From a Huge Selection! Games.com
Ad Formats	Text Ad (Instructions) (10)

Your Rating (11)

Ad Creative Flags	<input type="checkbox"/> Navigational Bullseye - guidelines (12) <input type="checkbox"/> Unexpected Porn or Download – guidelines <input type="checkbox"/> Foreign Language – guidelines <input type="checkbox"/> Awkward - guidelines <input type="checkbox"/> Competitive - guidelines <input type="checkbox"/> Geographic Off Target - guidelines
Ad Creative Rating	 <p> <input type="checkbox"/> Satisfaction Likely <input type="checkbox"/> Satisfaction Possible <input type="checkbox"/> Dissatisfaction Possible <input type="checkbox"/> Dissatisfaction Likely (13) </p>
Landing Page Flags	<input type="checkbox"/> Navigational Bullseye – guidelines (14) <input type="checkbox"/> Unexpected Porn or Download – guidelines <input type="checkbox"/> Foreign Language – guidelines <input type="checkbox"/> Error Did Not Load - guidelines <input type="checkbox"/> Competitive - guidelines <input type="checkbox"/> Geographic Off Target - guidelines <input type="checkbox"/> Product Not Found - guidelines <input type="checkbox"/> Affiliate/PPC - guidelines <input type="checkbox"/> Registration Form - guidelines
Landing Page Rating	 <p> <input type="checkbox"/> Satisfaction Likely <input type="checkbox"/> Satisfaction Possible <input type="checkbox"/> Dissatisfaction Possible <input type="checkbox"/> Dissatisfaction Likely (15) </p>
Overall Quality Rating	 <p> <input type="checkbox"/> Excellent <input type="checkbox"/> Fair <input type="checkbox"/> Poor <input type="checkbox"/> Terrible (16) </p>

Gut Reaction Rating	 (17)																					
Comment	<div style="border: 1px solid black; padding: 5px; min-height: 100px;"> (19) </div>	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 2px;">#ALERT#</td> <td style="padding: 2px;">Summon an Admin (18)</td> </tr> <tr> <td style="padding: 2px;">#DEADLOCK#</td> <td style="padding: 2px;">Rating Deadlock</td> </tr> <tr> <td style="padding: 2px;">#IG#</td> <td style="padding: 2px;">Inventory Glitch</td> </tr> <tr> <td style="padding: 2px;">#U#</td> <td style="padding: 2px;">Useless</td> </tr> <tr> <td style="padding: 2px;">#H#</td> <td style="padding: 2px;">Harmful</td> </tr> <tr> <td style="padding: 2px;">#MC#</td> <td style="padding: 2px;">Misleading Creative</td> </tr> <tr> <td style="padding: 2px;">#NT#</td> <td style="padding: 2px;">Biased or Not Trustworthy</td> </tr> <tr> <td style="padding: 2px;">#OQ?#</td> <td style="padding: 2px;">Low Confidence</td> </tr> <tr> <td style="padding: 2px;">#SEC#</td> <td style="padding: 2px;">Secondary Interpretation</td> </tr> <tr> <td style="padding: 2px;">#SPELL#</td> <td style="padding: 2px;">Spelling Correction</td> </tr> </table>	#ALERT#	Summon an Admin (18)	#DEADLOCK#	Rating Deadlock	#IG#	Inventory Glitch	#U#	Useless	#H#	Harmful	#MC#	Misleading Creative	#NT#	Biased or Not Trustworthy	#OQ?#	Low Confidence	#SEC#	Secondary Interpretation	#SPELL#	Spelling Correction
#ALERT#	Summon an Admin (18)																					
#DEADLOCK#	Rating Deadlock																					
#IG#	Inventory Glitch																					
#U#	Useless																					
#H#	Harmful																					
#MC#	Misleading Creative																					
#NT#	Biased or Not Trustworthy																					
#OQ?#	Low Confidence																					
#SEC#	Secondary Interpretation																					
#SPELL#	Spelling Correction																					

Cancel
(20)

Save Draft
(21)

Submit
(22)

The red numbers represent the following:

1. **ad rating => ad rating task**
Shows your location in the EWOQ system; in our screenshot, the display shows the path from **Ads Rating Home** (indicated here by "ad rating") to the current **Ad Rating Task** page.
2. **Ad Rating Task - [query]**
The heading of the **Ad Rating Task** page displays the query.
In our example, the text **Ad Rating Task - football games**, indicates that the query is [football games]
3. **release task**
This button releases the task back into the queue for another team member to work on. You will lose any work on the task you have done. It is possible to reacquire this task at a later date. A confirmation box will ask you if you really want to release yourself from the task before returning you to the Ads Rating Home page. The release task button is not available to you for tasks that have entered the deliberation stage.
4. **reject task**
This button permanently removes the task from your queue. It is rarely used. It is primarily used by team-members who do not feel comfortable evaluating tasks involving adult or pornographic content (such tasks are not very common). A confirmation box will ask you if you really want to reject the task before returning you to the Ads Rating Home page.
5. **Query and optional query research task button**
The query is shown above the "research this query" button. The button is provided if you do not fully understand the user's intent. The button opens up a separate browser window or tab and shows Google search results for the rating country and language for this task. Your understanding of the query is required; however, you only need to research the query if you don't understand the user intent.
6. **Language**
Indicates the language rating group. Each rating group has a Rating Language and a Rating Country. Some rating groups require you to use special instructions while rating tasks from that group. Administrators will tell you what groups you should work in and if your group requires special instructions beyond the regular guidelines. Most team members begin with access to the English-ads group. This group's rating language is English, and its rating country is the United States. If you were hired to work on a different language, administrators will add your rating language after you have completed your training with the English-ads group.

7. **Query Description and Due Date fields**
These fields are usually blank and can be ignored.
8. **visit landing page button**
This button is *inactive* until you choose an ad creative rating. Once you choose an ad creative rating, the button becomes active and you may click on it to view the landing page in a new window or tab to continue your analysis of the task.
9. **Ad Creative**
This is the advertisement that you must evaluate in relation to the query and according to the criteria described in the guidelines. Imagine that this ad appears on the search results page illustrated in screenshot 1.3.2 after the user enters their query.
10. **Ad Formats**
This line describes the type of ad format. Google displays different ad formats to its users: some ads may have maps, pictures, links to multiple products, etc. Some ad formats require you to do nothing more than read them before you choose your Ad Creative rating. Other ad formats require that you interact with the ad and consider flags specific to that format. Click on the [instructions](#) link to learn more about how to evaluate an ad format.
11. **Your Rating**
This section contains your ratings for the task. When you first acquire a task, you will have to choose all your scale ratings and determine which, if any, check boxes must be checked. Finally, there is a comment section, where you can comment about the task or insert comment codes, if applicable.
12. **Ad Creative Flags**
There are six labeled checkboxes in this section. Each term has a specific meaning defined in these guidelines. Clicking on the [guidelines](#) link for a check box will display a summary of the flag from the guidelines. Click on a box if you determine that the criteria are met for that flag.
13. **Ad Creative Rating slider**
This slider is divided into four ratings: Satisfaction Likely, Satisfaction Possible, Dissatisfaction Possible, Dissatisfaction Likely. Move the slider to the appropriate rating based on your evaluation of the ad creative.
14. **Landing Page Flags**
There are nine labeled checkboxes in this section. Each term has a specific meaning defined in these guidelines. Clicking on the [guidelines](#) link for a check box will display a summary of the flag from the guidelines. Click on a box if you determine that the criteria are met for that flag.
15. **Landing Page slider**
This slider is divided into four ratings: Satisfaction Likely, Satisfaction Possible, Dissatisfaction Possible, Dissatisfaction Likely. Move the slider to the appropriate rating based on your evaluation of the landing page.
16. **Overall Quality slider**
This slider is divided into four ratings: Excellent, Fair, Poor, Terrible. Move the slider to the appropriate rating based on your evaluation of the overall quality of the advertiser site.
17. **Gut Reaction slider**
This slider is divided into four ratings: Strongly Like, Moderately Like, Moderately Dislike, Strongly Dislike. Choose the appropriate rating based on your personal reaction to the advertising experience.
18. **Comment Code list**
This is a list of the most common comment codes used in ads quality evaluation. Comment codes are abbreviations used within ads quality evaluation to communicate with administrators and other team-members working on a task. Clicking on a code inserts it into the comment box. Each code has a specific meaning within ads quality evaluation. Section 5 describes all comment codes and their use.
19. **Comment box**
Enter comments or comment codes into this box.

20. **Cancel**

This button cancels any work you have done for the current Task and returns you to the Ads Rating Home page. Your current work is not saved.

21. **Save Draft**

This button saves a draft of your task and leaves the unfinished task in the Unrated or Partially Rated section of the Ads Rating Home page. You can return to the task later by clicking on it from the Ads Rating Home page.

22. **Submit**

This button is *inactive* until you have chosen ratings for ad creative, landing page, overall quality, and gut reaction. When you have fully evaluated the task, chosen ratings and flags, and left your comments, click the submit button to send it to EWOQ for evaluation. *Submitting a Task will remove it from your Task list. You will not be able to revisit or revise the Task unless the EWOQ system determines further input is needed on the task.*

Once you submit a task, the EWOQ system will automatically assign you a new task. Every 10 tasks, EWOQ returns you to the Ads Rating Home page to give you an opportunity to see if any tasks have moved to the Resolving or Held Open by Admin categories.

3.5 Reviewing and Re-rating Unresolved Tasks: Task Deliberation

Every task is acquired and rated by a group of raters, each working independently. When every rater has submitted his or her ratings, the EWOQ system compares your ratings and flags. Most of the time, raters more or less agree. The scores are recorded and analyzed, and no further action from you is required on the task.

When raters disagree with one another by a substantial margin, the task is returned to the raters involved for discussion and possible adjustment of initial ratings. These tasks first appear highlighted in yellow in the Resolving section of the Ad Rating Home page. When you see new tasks appear in this section, you must review them.

The process of reviewing tasks is called **deliberation**: the process of discussing all sides of an issue. There may be one or several issues to address with a task: the EWOQ system helps by highlighting in yellow the ratings and flags where disagreement is strongest. You and the other team members who rated the task are responsible for discussing it. A sample task in deliberation is examined below.

3.5.1 Ad Rating Task in Deliberation Screen Shot

ad rating → ad rating task jane.doe [ad rating – experimental – logout]

Ad Rating Task – football games

release task - reject task

Query	football games research this page (optional)
Language	English-ads
Query Description	None
Landing Page	visit landing page
Due Date	None
Ad Creative	Free Sports Games Play Free Sports Games at Games.com Choose From a Huge Selection! Games.com
Ad Formats	Text Ad (Instructions)

All Ratings (1)

Scores (2)

Rater	Last Modified	Ad Creative Rating	Landing Page Rating	Overall Quality Rating	Gut Reaction Rating
Rater 1	6/15/11 4:46 PM	Satisfaction Likely (100)	Satisfaction Likely (70)	Fair (40)	Hate It
Rater 2	6/15/11 4:35 PM	Dissatisfaction Likely (-57)	Satisfaction Likely (75)	Excellent (100)	Love It
Rater 3	6/15/11 4:49 PM	Satisfaction Possible (35)	Satisfaction Possible (22)	Excellent (100)	Love It
(Me) Rater 4	6/15/11 4:53 PM	Dissatisfaction Likely (-55)	Dissatisfaction Possible (-22)	Excellent (100)	Like It

Flags (3)


Rater	Last Modified	AC Flag Awkward Insertion	AC Flag Competitive	AC Flag Geo-Off Target	Landing Page Flag Competitive	Landing Page Flag Geo-Off Target	Landing Page Flag Product Not Found	Landing Page Flag Aff/PPC	Landing Page Flag Registration Form
Rater 1	6/15/11 4:46 PM								X
Rater 2	6/15/11 4:35 PM								
Rater 3	6/15/11 4:49 PM								
(Me) Rater 4	6/15/11 4:53 PM								

Comments on this Rating (4)

Comment	Rater	Timestamp
Site is ok, but they don't have any football games.	Rater 1	6/15/11 4:46 PM
User was probably not looking for online games, but actual live football games.	Me (Rater 4)	6/15/11 4:51 PM
Rater 1, your AC and LP seem pretty optimistic given what the user is looking for. Did you find football games?	Rater 3	6/15/11 4:59 PM

Your Rating (5)

Ad Creative Flags	<input type="checkbox"/> Navigational Bullseye - guidelines <input type="checkbox"/> Unexpected Porn or Download - guidelines <input type="checkbox"/> Foreign Language - guidelines <input type="checkbox"/> Awkward - guidelines <input type="checkbox"/> Competitive - guidelines <input type="checkbox"/> Geographic Off Target - guidelines
Ad Creative Rating	
Landing Page Flags	<input type="checkbox"/> Navigational Bullseye - guidelines <input type="checkbox"/> Unexpected Porn or Download - guidelines <input type="checkbox"/> Foreign Language - guidelines <input type="checkbox"/> Error Did Not Load - guidelines <input type="checkbox"/> Competitive - guidelines <input type="checkbox"/> Geographic Off Target - guidelines <input type="checkbox"/> Product Not Found - guidelines <input type="checkbox"/> Affiliate/PPC - guidelines <input type="checkbox"/> Registration Form - guidelines
Landing Page Rating	
Overall Quality Rating	

Gut Reaction Rating	 <div style="float: right; text-align: right;"> Love It Like It Dislike It Hate It </div>		
Comment	<div style="border: 1px solid black; height: 100px; width: 100%;"></div>	#ALERT# #DEADLOCK# #IG# #U# #H# #MC# #NT# #OQ?# #SEC# #SPELL#	Summon an Admin Rating Deadlock Inventory Glitch Useless Harmful Misleading Creative Biased or Not Trustworthy Low Confidence Secondary Interpretation Spelling Correction

The red numbers represent the following:

1. **All Ratings**
Shows the ratings, flags, and comments submitted by you and other team members with a "Last Modified" timestamp. Everyone participating in a task will stay anonymous. In fact, all raters are identified by "Rater" plus a number. Administrators will be shown as Administrator, and Moderators are shown as Moderator.
2. **Scores**
You will be able to see your initial four scale ratings/scores with the timestamp. In this example, your ratings are identified as **Me (Rater 4)**. Your rater number will change depending on when you acquired the task in relation to other team members. Ratings with the greatest difference in scores are highlighted in yellow.
3. **Flags**
You will be able to see your initial flags, with the timestamps. In this example, your flags are identified as **Me (Rater 4)**. Flags in which there is substantial disagreement are highlighted in yellow.
4. **Comments on this Rating**
Your initial comments, if any, for this task. All comments include rater and timestamp information. As you and other team-members enter comments, they are posted in this box. Most recent comments appear at the bottom of the box. Occasionally, you will encounter administrator or moderator comments, highlighted in yellow. Some rater comments will appear crossed out, indicating it has been reviewed by an administrator.
5. **Your Rating**
This section lists your current ratings and flags. The EWOQ system highlights in yellow any rating or flag where there are big differences between team members' scores. If you want to make changes to your ratings or flags, make them here. You should also leave a comment explaining your ratings, changes to your ratings, how you're interpreting the query, or how you're interpreting or applying the guidelines.

When you have reviewed the task, made any changes to ratings or flags, and written your comment, submit your changes to the EWOQ system. The Cancel and Save Draft buttons work the same way and serve the same function as they do when you first evaluate a task.

3.5.2 The Purpose of Task Deliberation

The purpose of task deliberation is to discuss the various aspects of a task where raters disagree over their ratings, comment codes, query interpretation, or application of the guidelines. There are three equally acceptable conclusions to any task in deliberation:

- **Resolution:** A task is resolved when one or more raters change their ratings and/or flags so that they are in agreement with other raters.
- **Deadlock:** One or more raters decide that, after a full discussion, resolution is not possible or appropriate. See Section 5.1.5 for more information on the deadlock option.
- **Closure:** Administrators occasionally choose to close tasks for reasons other than resolution or deadlock.

Task deliberation is a discussion among all task participants. You are individually responsible for discussing your ratings, and you are welcome to ask other raters to discuss their ratings as well. Tasks usually enter deliberation for the following reasons:

- One or more raters entered incorrect scores/flags because of a mouse slip or other error.
- Raters are seeing different landing pages or experiencing the advertiser's site differently.
- One or more raters noticed something about the task that others missed.
- Some raters are rating the task based on a misunderstanding of the guidelines or are interpreting the guidelines differently from others.
- Real substantial differences of opinion exist about whether or not the ad or the page represent a positive or negative experience for the user entering the query, or whether or not the advertiser's site represents a positive or negative experience for any user.

When you encounter a task in deliberation, **all** team members must collectively determine the foundation of the disagreement. Double-check that your own ratings are appropriate, and try to understand why others rated the way they did.

Task deliberation is a time to diagnose the outstanding issues you and your colleagues are experiencing. When you first visit a Resolving task, we recommend you do the following:

- Don't initially assume someone is wrong. Try to diagnose the task based on the range of ratings, flags, and comment codes everyone has used.
- Carefully review your own ratings, flags and comment codes.
- Write a comment that speaks to the facts: what you're seeing, how you're interpreting the query, or something you saw that you think others may have missed.
- Discuss the task or ask questions in a polite and professional manner.
- Be patient: not everyone has the same work schedule, and not everyone can respond immediately. Deliberation usually takes time.

Everyone makes a mistake occasionally, and this is to be expected. If a task enters deliberation and you think you made a mistake, go ahead and correct the mistake and move on. If you find yourself regularly making mistakes, forgetting flags, or comment codes, you are probably rating carelessly. Slow down and try to evaluate your tasks more thoroughly and accurately.

If you rated a site positively because you didn't realize the site was a scam, the appropriate course of action is to choose the ratings for scams. Downgrading to Dissatisfaction Possible or Poor is just as wrong as Satisfaction Possible or Fair if the correct rating is Dissatisfaction Likely or Terrible. Deliberation is no place for egos or "saving face". Stick to the facts, and always try to choose the most accurate ratings. Don't be afraid to ask questions, and don't be embarrassed if you initially missed something that seems obvious when viewing it again. Just choose the right ratings and move on.

If you find that you made a mistake in your ratings, flags, comment codes, query interpretation, or application of the guidelines it is extremely important that you update your ratings and flags according to how the guidelines tell you to rate.

Please note that you are required to discuss ratings and flags highlighted in yellow, but you are still responsible for ensuring all your ratings and flags are correct. When reviewing a task please

quickly review all your ratings and flags to make sure they're correct, even those that are not highlighted in yellow by the EWOQ system.

3.5.3 Administrator and Moderator Comments

Administrators and moderators also review tasks in deliberation. Both may comment on tasks and provide guidance or direct raters to expand on a discussion.

Administrators work at Google. They review and monitor every deliberation, but do not always comment: the raters assigned to a task generally have all the resources they need to deliberate a task among themselves.

Administrators may intervene under the following circumstances:

- Raters are seeing different landing pages or sites.
- The task is defective or inappropriate for ads quality evaluation.
- They are asked by a rater to review deadlock requests.
- Someone is not participating fully or appropriately in deliberation.
- Raters have a fundamental disagreement about how to interpret or apply the guidelines and ask for guidance.

Administrators will usually address all raters on a task, though sometimes they will address individuals. In any case, you should study administrator comments and respond appropriately. Sometimes, administrators will move Tasks in the Resolving section to the Held Open by Admin section. Regardless of where you see a task, you should always review and consider administrator comments. The Held Open by Admin section is discussed in Section 3.7, below.

Moderators are senior ads raters. When they're not rating tasks themselves, they review tasks that have been alerted, deadlocked, as well as tasks that have been in discussion for a long time. Moderators can only comment: they do rate the task. Moderators can answer common questions about the guidelines. Moderators do not examine all tasks: raters are expected to handle discussion themselves for the most part.

3.6 Managing Open Tasks

Once you begin rating tasks, your Ads Rating Home page displays your current list of Acquired Tasks. You may visit the list at any time by selecting the "ad rating" link in the upper left and right-hand corners of web pages in EWOQ.

3.6.1 Task Management: Ad Rating Home Screen Shot

ad rating jane.doe

[[ad rating](#) – [experimental](#) – [logout](#)]

Ads Rating Home

[acquire next ad rating task](#) - [history](#) - [guidelines](#) - [hub](#)

Acquired Tasks

Unrated or Partially Rated Tasks

Query	Last Viewed by Me	Last Modified by Me	Last Modified by Someone Else	Status	Modified by Someone Else since Last Viewed
music videos	Jun-13-2011 12:28	Jun-13-2011 12:29	Jun-13-2011 13:01	Unresolved	Yes

Held Open by Admin

No tasks in this category...

Resolving

Query	Last Viewed by Me	Last Modified by Me	Last Modified by Someone Else	Status	Modified by Someone Else since Last Viewed
pearls	Jun-13-2011 12:28	Jun-13-2011 12:29	Jun-13-2011 13:01	Unresolved	Yes

Resolved but not Closed

Query	Last Viewed by Me	Last Modified by Me	Last Modified by Someone Else	Status	Modified by Someone Else since Last Viewed
honda accord	Jun-13-2011 12:28	Jun-13-2011 12:29	Jun-13-2011 13:01	Resolved	Yes

In the screenshot, above, there are tasks in three of the four categories. As you work you will find that you regularly have several tasks in each category.

3.7 Prioritizing your Tasks

Your primary responsibility as an ads quality evaluator is to acquire, evaluate, and submit rating tasks. You are responsible for evaluating and submitting tasks that you acquire and discussing any task that appears in the Resolving or Held Open by Admin categories.

Prioritization of Tasks in your queue - **Acquired Tasks**

Highest Priority: Unrated or Partially Rated categories

High Priority: Held Open by Admin and Resolving categories

Optional: Resolved but not Closed category

The four Acquired Tasks sections are listed in order of how we expect you to prioritize your work. Unrated or Partially Rated Tasks are the highest priority because your primary responsibility is to

acquire and evaluate new tasks. When you acquire tasks, you're telling the EWOQ system you want to rate them. Whenever you want to work, you should acquire, rate, and submit new tasks.

It is also very important to review tasks that you have already evaluated that are Resolving or Held Open by Admin. Everyone assigned to a task is responsible for reviewing all tasks in these sections. You may or may not need to make a change to your rating, but you must always review and discuss your rating decisions. If you have already provided input and responded to input by the administrator or other team members, there is no need to write further comments unless you think your input can help the team with their deliberations. If you've reviewed all your Resolving and Held Open by Admin tasks please return to evaluating new tasks.

All Held Open by Admin Tasks start as Resolving Tasks but are transferred by administrators so the team members assigned to the task know that the task requires their immediate attention. A task is usually Held Open by Admin because the administrator wants a fuller discussion or wants to speed up deliberation before the task is closed. When a task is moved to the Held Open by Admin section, you should review the task as soon as possible. Administrators will usually post a comment explaining their concern. If more input is needed by you, leave a comment and change ratings or flags if necessary. If you think your existing comments already address the concerns raised by the administrator and you feel comfortable with your ratings, an additional comment is optional. If you have reviewed all tasks in the Held Open by Admin and Resolving sections, you should return to your primary responsibility: acquiring, evaluating and submitting new tasks.

When a task is completely resolved, it is moved to the Resolved but not Closed section. You can still access tasks in this section, and you can still post comments or make changes to your ratings and flags. You are not required to make any changes to tasks in this section unless your ratings or flags are incorrect and need correcting. We include tasks here so you can see how tasks resolved before we remove them from your queue. You don't have to perform any actions to remove tasks from the Resolved but not Closed section: administrators will take care of them.

3.7.1 Tools to Help Manage Your Work

Sorting your tasks: The Ad Rating Task page (Screen Shot 3.6.1) provides you with tools to help you manage your workflow. The categories of Acquired Tasks will either be empty (section will say "No tasks in this category...") or there will be a list of tasks in that category. Tasks listed in any of the categories can be sorted by clicking on any column title. If you want to sort your Resolving tasks in alphabetical order, click on the column title you want to sort by. Clicking on the column title again sorts the tasks in reverse order. For example, if you want to sort tasks in order of when you last viewed them, click on the "Last Viewed by Me" column.

The final column (Modified by Someone Else Since Last Viewed) helps raters resolve tasks in deliberation more quickly. If a task in this column says "yes" it means someone else (including an administrator) has changed a rating or added a comment since you last viewed it. View the task to see what comment was posted or what ratings or flags have been changed.

History Function: Clicking on the history link on the Ad Rating Home page (Screen Shot 3.3.1) permits you to review the tasks you've rated. Note: we do not recommend that you use the history function to fill out your timesheet as the timestamps don't represent the moment you began rating but when you submitted a task. History also does not include experimental tasks you work on.

3.8 Task Deliberation Etiquette

The following are guidelines for effective task deliberation in EWOQ.

1. Always share relevant background information (reasons, explanations, etc.) when stating your opinion. Indicate your source of information whenever possible. If you come across an important website in your research, please give its full URL.
2. The people reading your comments need to know what you're referring to. If you're referring to the ad creative, begin by saying you're talking about the AC. If you're talking about the overall quality rating, start out your comment by referring to OQ.
3. When commenting on a query, *describe your interpretation of the user intent*. This is important for ambiguous or poorly phrased queries. You may discuss whether you feel the query is a do, know, or go query, what the dominant interpretation is, etc.
4. Please do not use abbreviations except for those listed below.
Exception: To save space and time, all comment codes (i.e. #U# and #IG#), and the following abbreviations for ratings and flags may be used:

AC	Ad Creative	Comp	Competitive
LP	Landing Page	GOT	Geographic Off Target
OQ	Overall Quality	Aff	Affiliate
E/DNL	Error/Did Not Load	Aff/PPC	Affiliate/Pay Per Click
UP/UD	Unexpected Porn/ Unexpected Download	RegForm or RF	Registration Form
SL	Satisfaction Likely	DL	Dissatisfaction Likely
SP	Satisfaction Possible	DP	Dissatisfaction Possible
FL	Foreign Language		

5. Please refrain from using message board lingo (IMO, FWIW, AFAIK, etc.).
6. Write concisely. Do not make unnecessary comments such as "Oh, I see your point" or "Sorry, I missed that". But do write enough to explain yourself clearly to other raters who might not have your background or expertise.
7. Sometimes the most efficient way to make your point is to quote guidelines. Please be very specific about how the information you quote relates to the situation at hand. When quoting from the guidelines, please include the page or section number.
8. Be ready to state the reason for assigning a Comment Code if the task enters deliberation. For example, if you initially entered a comment code when you rated a task, these would good followup comments:
 - "nothing but a list of ads" (if you had entered #U#)
 - "user is likely to become the target of email spam" (if you had entered #H#)
9. Brief comments to confirm your rating in the resolving stage are always appreciated:
 - "Reviewed. AC still deserves DL for me."
 - "Confirming no Comp flag: advertiser is retailer, not manufacturer sought in query."
10. If another rater or administrator asks you to comment, please do so.
11. Remember that raters have different work schedules: don't assume someone is ignoring your comments just because they haven't responded after a few hours. Some deliberations can take a few days to resolve. Be patient!

Things to avoid in Task Deliberation:

1. Non-comments like "oops", ".", or "ok".
2. Comments like "d/g", "changing", "u/g my LP". Such comments don't tell us *why you are changing your ratings*. If you've already explained yourself, then a brief comment is ok.
3. Rude or unprofessional comments.
4. Except for comment codes, the use of ALL CAPS. Some people consider it the equivalent of yelling.

4.0 Rating a Task

The following section describes the scale ratings and flags you use for each task you evaluate. Please read this section very carefully. You should refer to this section regularly as you evaluate tasks and discuss your ratings with other team members.

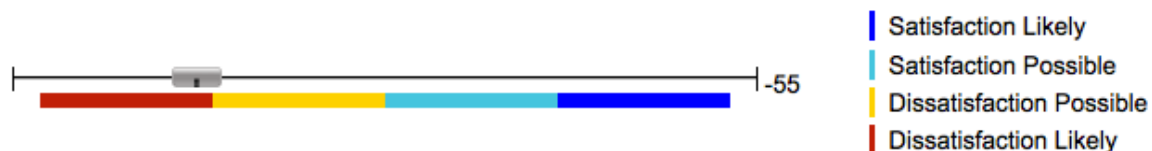
Remember, for Ad Creative and Landing Page ratings, you must already have researched the query. Your AC and LP ratings are based on your interpretation of the query and how well (or poorly) the advertiser responds to that query interpretation.

4.0.1 Using the Slider Bar

The slider bar is used for your Ad Creative, Landing Page, Overall Quality and Gut Reaction scale ratings. You must select a value for each rating before you can submit a task. Each scale rating is divided into four sections, each of which corresponds to a specific score:

- **Ad Creative** and **Landing Page**: chose one of four scores ranging from Dissatisfaction Likely to Satisfaction Likely.
- **Overall Quality**: choose one of four scores ranging from Terrible to Excellent.
- **Gut Reaction**: choose one of four scores ranging from Hate It (Strongly Dislike) to Love It (Strongly Like).

If you want to choose a score, move the slider anywhere within the range of that score. Start your slider in the middle of the score. If you feel it is a strong or weak example of that rating, move it to the left or right. For example, the rater below has chosen the Dissatisfaction Likely score for her Ad Creative Rating, but she moved the slider close to the adjacent Dissatisfaction Possible score. Her score communicates that this Ad Creative meets the criteria of Dissatisfaction Likely, but only barely so. She feels it almost meets the criteria of Dissatisfaction Possible.



Where you place your slider within a score is as important as which score you choose. When a task enters deliberation, discuss which score should be used with other raters. You may also need to talk about where the most appropriate place to put the slider within that score is.

You may find it difficult to decide between a positive and negative rating or a strong and a moderate rating. Keep in mind that if you're having a hard time deciding between two ratings, other raters may be having the same difficulty. Be bold in your ratings and your analysis, and always keep the user experience in mind: the more accurately you evaluate the user experience the more we learn about what ads are useful to users and what ads aren't. And remember, if your task enters the deliberation stage, you will have the opportunity to share your thinking on the task, review the evidence you found, and discuss *why* it was hard to make a decision.

4.1 Ad Creative

The ad creative is an advertisement written to appeal to the user entering the query. The advertiser hopes the user will click on the ad. Clicking on the ad takes the user to the advertiser website.

4.1.1 Ad Creative Overview - What You're Doing

When you rate the ad creative, you tell us how well the advertisement matches the User Intent. In general, a **satisfactory** ad either offers something that will satisfy the user intent, or something the user is likely to find interesting and relevant. A **dissatisfactory** ad offers something that doesn't satisfy the user intent, or there is little reason to think the user will be interested in what the advertiser offers. Some dissatisfactory ads are related to the words of the query, but not the user intent. Others simply have nothing at all to do with the query or the user intent.

Sometimes an ad will offer pornography or sexually-related material to a user who didn't ask for it. Those are **very bad** ads. They give the user a bad experience and make the user less likely to use Google. Sometimes an ad looks like it is trying to cheat or con the user. Those are also **terrible** ads. Google should never show ads like that.

Finally, sometimes ads will be in a language the user probably doesn't speak. Those are also **very bad** ads. They completely waste the user's time, and make Google look like it doesn't know what language a person uses.

This is what Google wants to do:

- **Show ads relevant to what the user wants**
- **Show ads the user will consider appealing**
- **Don't show ads unless there's good reason to think they're relevant to what the user wants**
- **Don't show ads that will offend the user or lead the user to harm**

Some users are interested in ads. Some users ignore them completely. Some users hate them. You can't tell from looking at the query whether the user will actually click on the ad, and we don't expect you to guess. Even if the user is someone who would never click on ads, we want the ads that we show to be the best ones possible. Your objective analysis of the user's query and the ad helps us determine when whether we're doing this or not.

The **best** kind of ads offer exactly what the user is looking for. If the User Intent is to buy a widget, the best kind of ad is an ad for widgets or an ad for a place that offers widgets.

The **second best** kind of ad is one that doesn't satisfy the User Intent directly, but is very relevant to the User Intent, and offers something that the user might **also** be interested in looking at. For example, a user who is looking for [cheap flights to Paris] might also be interested in inexpensive hotels in Paris. An ad for inexpensive hotels in Paris is a good ad to show! It doesn't offer what the user was actually looking for, but it makes sense to see if the user might find it useful. If the idea behind an ad seems to be "If you're interested in that, then you're probably also interested in this"--and you think it's true--it's a good ad. Unless there's good reason to think the ad is relevant to the user intent, there's not much point in showing it. We still need to learn from you why it isn't a good ad for the user.

If the ad is only generally related to the User Intent, but there's no reason to think the user would be interested in what the advertiser offers, it's only **mildly bad**. We still don't want to show it, but it is not a disaster if we do.

If the ad is completely unrelated to the User Intent, it's **very bad**. It's also really bad if the ad "misunderstands" words in the query--for example, if the query is [trips to java] and the ad is related to the Java programming language. Ads like that make Google look stupid, and make the user feel like Google ads aren't worth looking at.

It's also **very bad** if the ad looks like a scam. "Too good to be true" offers and obvious lies are a really bad experience, even if they don't click on the ad. They may fool some people, but if an ad looks like a scam to you, you should consider it a very bad ad. Google doesn't want users to be victims of scams, especially when the scam is offered by an advertiser.

Ads in the wrong language are a complete waste of the user's time. Ads that offer pornography or sexual services when that's not what the user is looking for are an **extremely bad** experience. Both of these things are so bad that you will use special flags just for them.

When you evaluate the relevance and usefulness of an ad for the user, you first check any flags, if applicable, and then choose your scale rating. The flags and rating chart for Ad Creative are found in the next two sections.

4.1.2 Ad Creative Flag Overview

Every Ad Creative has at least five flags (checkboxes) you must consider before choosing your scale rating. You will occasionally see new ad formats, and some of these ad formats have additional flags for you to consider. All flags have a link to the guidelines if you need help determining if the flag should be checked.

4.1.3 Ad Creative Flag Reference

Navigational Bullseye

If a query is navigational, and the creative looks like it will point to the site the user wanted, use the **Navigational Bullseye** flag. Selecting this flag automatically moves your AC slider to Satisfaction Likely (100). If you decide the AC is not a Navigational Bullseye, uncheck the box.

Unexpected Porn or Download

If the ad creative offers pornography, sexual services, or sex dating, use the **Unexpected Porn or Download** flag, whatever the query is. Selecting this flag automatically moves your AC slider to Dissatisfaction Likely (-100). If you decide the AC is not an ad for pornography, uncheck the box.

If you think the dominant interpretation of a query is a search for pornography, sexual services, or sex dating, please alert administrators so the task may be reviewed and removed.

An ad creative can't download anything unexpectedly: we call it this to keep the name consistent between ad creative and landing page ratings.

Foreign Language

If the ad creative is not in your rating language, use the **Foreign Language** flag even if you personally can understand the language. If you evaluate tasks in multiple languages, confirm the rating language of the task first and only use the flag if the page is in the wrong language for the specific task you are working on. Selecting this flag automatically moves all sliders to -100, since we assume a task in the wrong language cannot be evaluated. If you decide the AC is not an ad in a foreign language, uncheck the box.

If the landing page contains a mix of languages, including your rating language, use the Foreign Language Flag if the user would be unable to determine how relevant or useful the ad creative is for his or her query.

Awkward Insertion

If the words in the ad creative are out of order or awkwardly inserted, use the Awkward flag. If the wording of the ad is unintentionally funny, bizarre, or offensive, use the Awkward flag.

Ad Creatives should make sense and sound good if you were to read them out loud. Sometimes, ads are put together in a way that makes them sound strange, funny, or even offensive. Sometimes they don't make any sense at all. You will use the Awkward Insertion Flag to tell us about ads like this. In general, you should use the Awkward Insertion Flag whenever:

- The words of an ad are out of order
- The words of an ad don't make any sense
- The wording of an ad is unintentionally funny, bizarre, or offensive

There are a couple of ways that creatives can turn out badly. One of the most common involves keyword insertion. Advertisers sometimes provide creatives that have empty spaces in them, and these spaces are filled in by words from the query or words that are related to the creative. This is called **keyword insertion**. Some advertisers have the ability to include some of the user's search terms in their ad creative. When this works well, the ad is better tailored to the user intent. When it works poorly, however, it can result in strange or even inappropriate ads. The Awkward flag identifies situations where keyword insertion does not work well.

There are many other ways that ads can deserve Awkward Insertion. Let your ear be your guide. If an ad sounds weird to you, mark it Awkward Insertion, even if you can't put into words exactly what's wrong with it. Also keep in mind that you don't need to check grammar the way a grade-school teacher would. Advertisers regularly omit a preposition or article. Don't use the Awkward Insertion Flag unless the omission results in an ad where regular users think an ad sounds strange or bad.

Competitive

Use the Competitive flag when the ad creative of an advertiser suggests direct competition with a product, brand, or company in the user's query. The advertiser must actually compete (or offer a competing product or brand) in order to be competitive. For example, if a user is looking for [nike air], a general ad creative by Foot Locker is not competitive since Foot Locker sells athletic shoes. An ad by Adidas, however, is competitive. Adidas and Nike compete with each other. An ad by a shoe retailer that only mentions Adidas would only be competitive if you thought the advertiser probably didn't sell products by Nike.

The ad creative alone may not always be enough to determine whether there is potential competition. If you are unfamiliar with the product, brand, or company specified by the user entering the query, research it. Query research involves not only an understanding of user intent but also knowledge of the queried terms entered by the user. A user knows what he or she is looking for. If the user would think the advertiser is a competitor (or offering a competing product or brand, but not the queried brand) then you should use the flag.

General queries that do not mention specific products, brands, or companies never receive the Competitive flag. In general, reserve Competitive for fairly obvious cases. If it seems like a little bit of a stretch in a particular case, go ahead and use it. If it seems like a big stretch, don't.

Geographic Off Target

The Geographic Off Target (GOT) flag is used when the advertiser offers a product or service that is incompatible for the user entering the query because of geographic location. There are two situations in which the flag is used:

GOT - Situation One: Wrong Location

A prerequisite to using this flag is that the user must mention a city, state, region, or country in the query. General queries that do not mention a specific location do not get the GOT Flag. The ad creative must specify a product, service, or location that is DIFFERENT and INCOMPATIBLE with what the user is looking for in the query.

Incompatibility based on distance from the queried location will vary depending on the query. A user searching for [plumber novato] requires the plumber be within easy driving distance of the user's location. A user searching for [vacation kauai] might be interested in locations other than the queried one. A resort in neighboring island, Oahu, is over a hundred kilometers away, but it might be of interest to a user doing vacation research. An advertiser that provides nationwide shipping is compatible with a user searching for [pool cues waltham ma]: yes, the user specifies a specific location but the advertiser can reasonably ship products to the user. This is not GOT.

Geographic overbroad or overspecific advertisements do not receive the GOT Flag. For example, an ad offering hotel listings throughout France is compatible for the user query [hotels marseilles], and an ad for hotels in Marseilles is compatible for the user query [hotels france]. The GOT Flag is not used with geographically non-specific locations. For example, an ad creative for a San Francisco hotel offered for the query [countryside bed and breakfast] is not GOT. While there is no countryside in San Francisco, the user has not indicated a specific location. This may not be a good match for the query, but it is not GOT.

GOT - Situation Two: Rating Country Error

Every task you evaluate has a rating country. That is, the query comes from a user within the borders of the rating country, and we expect advertisers to target these users. If an ad creative targets users from a DIFFERENT country, it should be rated GOT unless something in the query makes you think that the products or services offered by the advertiser would be desirable to users in your rating country.

Despite being only a few lines of text, the ad creative will sometimes provide clues as to what country the advertiser is targeting, or where the advertiser is based. Some clues are:

The visible URL in some ads tell you what country the advertiser is based. This is called a top-level domain, and can either be non-specific (like .com or .org) or can refer to specific countries (like .be and .at, for Belgium and Austria, respectively). Non-specific URLs usually refer to organizations in the United States of America, but not always. Country-specific URLs usually refer to organizations within that country. There are some exceptions: for example, .tv is the top-level domain for Tuvalu; however, advertisers who use the .tv suffix rarely target users from Tuvalu.

Listing a currency (such as £ for British pounds or \$ for United States dollars, or AU\$ for Australian dollars) can be a clue as to what national audience the advertiser is targeting, but this alone isn't a good reason to use GOT. A user may either seek results from another country (and thus another country's advertisers), or an advertiser from another country may have good reason to advertise to the user. Use your best judgment to sort it out. For example, a user in the United States seeking [rare antiquarian books] might benefit from an advertiser in the United Kingdom; however, it is unlikely the same user seeking [mouse pads] will benefit in the same way. If an item is difficult to find, a user might search more broadly for it.

If an advertiser specifically targets or addresses users in your rating country, do not consider the ad creative GOT even if buying the product or service from abroad seems unlikely.

Additional Ad Creative Flags

You may find Ad Creative flags in addition to these five. Additional flags are included with new ad formats. You must evaluate the Ad Creative for every flag listed. The “guidelines” link next to each flag always provides details to help you determine if the flag is applicable or not.

4.1.4 Ad Creative Rating Reference Chart

Consult the chart below after you have evaluated the user query and determined the User Intent. Select an ad creative rating that best corresponds with expected user satisfaction. Review the Positive User Experience and Negative User Experience sections of the chart to decide if the AC is more likely to be a positive or negative experience for the user who entered the query. Then review the criteria for each of the individual ratings to determine the most appropriate AC rating.

With your ad creative analysis, you are making an educated guess as to whether or not the advertiser will satisfy the user based solely on the visible ad creative. You choose your rating without viewing the landing page (in fact, the “visit landing page” button will not be active **until** you select an ad creative rating).

Once you have chosen a rating, decide *where* within that rating the slider should be placed. If you believe the AC strongly meets the criteria of a rating choose a bigger number (i.e. a bigger negative number for negative user experience ads and a bigger positive number for positive user experience advertiser ad creatives). While technically possible to choose a “0” rating, we expect you to make a conclusion, either positive or negative.

Ad Creative – This chart helps you determine your AC rating, first by helping you consider what makes up a positive or negative user experience, and then by helping you decide which of the four ratings to use: Dissatisfaction Likely, Dissatisfaction Possible, Satisfaction Possible, Satisfaction Likely.

Focus on the user query: do you think that the user will be satisfied if s/he clicks on the advertiser AC? Do you think User Intent will be met by clicking on this ad?

Negative User Experience

Use a negative rating if the ad meets any of the general criteria below:

- The ad is probably not going to lead to a satisfactory experience.
- The ad is poorly formatted or difficult to determine whether or if the user intent will be met by this advertiser.
- The advertiser is not a good match for what the user is looking for: it is not clear how the user intent will be satisfied.
- The ad meets any of the specific criteria below:

Positive User Experience

Use a positive rating if the ad meets any of the criteria below.

- The ad probably is going to lead to a satisfactory experience.
- The ad is clear and easy to understand, and it is easy to determine how the advertiser will meet the user intent.
- The ad directly addresses the needs of the user.
- The ad meets any of the specific criteria below:

Individual criteria are listed below each column. An ad creative need only meet one criterion in a column for that rating to be a possible choice. If criteria from multiple columns apply to your ad creative, use your best judgment and be prepared to discuss the reasons for your decision with other raters.

Dissatisfaction Likely	Dissatisfaction Possible	Satisfaction Possible	Satisfaction Likely
<p>The ad has no connection to the User Intent.</p> <p>The ad looks like it could lead to a page or web site that may harm the user.</p> <p>The ad makes no sense.</p> <p>There is no clear way the user intent can be met if the user were to click on the ad.</p> <p>The advertiser offers pornography or other clearly adult content, and the user did not seek such content. (UP/UD)</p> <p>The ad is not in your rating language. (FL)</p> <p>The user query indicates a search for something in a specific location, and the advertiser offers a service or product in a different and incompatible location. (GOT)</p> <p>The advertiser is in a different country, and there is nothing in the user query that makes you think the user would be interested in clicking on an ad from a different country. (GOT)</p>	<p>The ad is confusing enough that you're not confident the user will have a satisfactory experience with this advertiser.</p> <p>This ad has only a general or marginal connection to the User Intent.</p> <p>The ad is poorly written, and it is difficult to determine if clicking on it will lead to a satisfactory experience.</p> <p>You can identify specific elements of the ad that lead you to conclude the user is not likely to find a product or service relevant to what the user is looking for.</p>	<p>Clicking on this ad probably isn't going to lead the user to the exact thing s/he wants, but it will lead to an appealing substitute.</p> <p>The user intent won't be satisfied by clicking on this ad, but the user will find other things that will be appealing (i.e. directly related products or services).</p> <p>The advertiser is a competing service or offers a competing product to what the user seeks, and you expect the user might be interested in what the advertiser offers.</p>	<p>Clicking on this ad probably is going to lead the user to what s/he wants.</p> <p>The user wants to buy a product or service, and clicking on this ad will lead the user to a page where a purchase can be made.</p> <p>The user seeks information, and clicking on this ad will lead the user to a page where this information can be found.</p> <p>The user seeks a specific web page. The ad looks like it will lead the user to that page (Navigational Bullseye)</p> <p>The user seeks a specific web site. The ad looks like it will lead the user to that web site. (Navigational Bullseye)</p>
Dissatisfaction Likely	Dissatisfaction Possible	Satisfaction Possible	Satisfaction Likely

4.2 Landing Page

When a user clicks on an ad creative, we send the user to the advertiser landing page. This is the first page of the advertiser site the user sees.

4.2.1 Landing Page Overview - What You're Doing

When a user clicks on an ad, the user is sent to the advertiser's landing page. When you give your landing page rating, you are telling us how well that page matches the User Intent. When you **rate the landing page**, you never consider your ad creative rating or the ad creative. They are two separate ratings.

There is an important difference between the *landing page* and the *landing site*. Both are important for your landing page rating, but the *landing page* is more important. The *landing page* is the actual page that the user sees first. The *landing site* is the entire website of which that page is a part. Often, you will need to look at more of the site than the *landing page* to determine your rating. Always keep in mind, though, that the landing page is what the user will see first. If the landing page is bad enough, the user will never bother to look at the rest of the site, even if it's wonderful!

As with ad creative, always remember that you are judging how relevant the page is to the User Intent, not just the words of the query.

The criteria are very similar to those for ad creative. The best kinds of landing pages **satisfy** the user intent. If the user wants to buy something, this is a page where they can buy that thing. If the user wants information, this is a page where they can find that information. If the user wants a particular web page, this **is** that web page.

The next best kind of landing page offer something **closely related** to what the user is looking for, and likely to be appealing to that user, even if they don't actually offer what the user desired. For example, if a user is looking for reviews for a particular product, a page where you can buy that product doesn't actually satisfy the user intent, but it's closely related and likely to be of interest to the user. Similarly, if the user is looking for cheap flights to Paris, a page offering cheap hotels in Paris isn't what the user was looking for, but there is good reason to think that the user might also find it useful.

Anything that doesn't fall into one of those two categories is a bad landing page and should be rated accordingly. Keep in mind that a page might be wonderful in general, but bad for a particular query. For example, www.amazon.com is a great site, but it's useless if you're trying to buy a tractor. We don't want to waste users' time by sending them to advertiser pages they won't find relevant or useful.

Pages that are somewhat related to the user intent, but not so relevant that you think that user has a good chance of being interested in it, are mildly bad. Google shouldn't send users to those pages, but it's not disastrous if they do.

Pages that have nothing to do with the user intent (for example, that have no connection to the query, or are relevant to some words of the query but the advertiser has misunderstood what the user was looking for) are very bad results. Being sent to a page like this is always a waste of the user's time.

Keep in mind that your landing page rating is based on how well the advertiser responds to and fulfills the User Intent. You are verifying whether the advertiser actually does satisfy the user intent. In addition, when you are looking at the landing page, you need to judge whether the page

or site actually is a scam or otherwise harmful to users.

4.2.2 Landing Page Flag Overview

Every Landing Page has nine flags you must consider before choosing your scale rating. Each flag is explained below. There is also a guidelines help link adjacent to each flag on the ad rating page if you need help determining if the flag should be checked.

4.2.3 Landing Page Flag Reference

Navigational Bullseye

If a query is navigational, and the landing page is the exact page the user wanted, use the **Navigational Bullseye** flag. If a query is navigational, and the landing site is the exact site the user wanted, use this flag. Selecting this flag automatically moves your LP slider to Satisfaction Likely (100). If you decide the LP is not a Navigational Bullseye, uncheck the box.

Unexpected Porn

If the landing page offers pornography, sexual services, sex dating, or other clearly adult-oriented topic, use the **Unexpected Porn or Download** flag, whatever the query is. Selecting this flag automatically moves your LP slider to Dissatisfaction Likely (-100). If you decide the LP doesn't meet these criteria, uncheck the box.

If you think the dominant interpretation of a query is a search for pornography, sexual services, or sex dating, please alert administrators so the task may be reviewed and removed.

Unexpected Download

An advertiser that attempts to download a file to the user's computer without his or her informed consent is unexpected and a very negative user experience. Informed consent means the advertiser explains to the user what is being downloaded, and the user has the option (i.e. by clicking on a clearly labeled button) to choose to download the file or not. A button that doesn't clearly indicate a download will be initiated by clicking on it is unexpected and a very negative user experience. An advertiser that tries to trick the user into downloading a file (i.e. by hiding the details in fine print) is a very negative user experience.

If you conduct web research and determine the file download offered by an advertiser is a threat to users (i.e. it is malware, or it will harm the user's computer), it is a very negative experience and the Unexpected Download flag must be used. If another rater researched it and chose this rating, other raters are required to independently research the rater's claim.

Foreign Language

If the landing page is not in your rating language, use the **Foreign Language** flag even if you personally can understand the language. If you evaluate tasks in multiple languages, confirm the rating language of the task first and only use the flag if the page is in the wrong language for the specific task you are working on. Selecting this flag automatically moves all sliders to -100, since we assume a task in the wrong language cannot be evaluated. If you decide the LP is not in a foreign language, uncheck the box.

If the landing page contains a mix of languages, including your rating language, use the **Foreign Language** flag only if a user in your rating language would be unable to determine how relevant or useful the landing page is for his or her query.

Don't use the **Foreign Language** flag if the landing page is in the wrong language, but it is easy to display the correct language (i.e. via a language link or national flag icon), don't use the flag. Keep in mind that a landing page in the wrong language is not a positive experience, even if you can switch to the correct language.

Error Did Not Load

If you click on the "visit landing page" button and are unable to evaluate the advertiser landing page because it does not load, click the **Error Did Not Load** box. This includes blank pages with no text, server error pages (which may be a generic page provided by your browser or the advertiser's server).

If you encounter an expired certificate warning on an advertiser landing page, and your browser warns you about visiting the site, it is generally safe to visit and evaluate the site. Don't use the EDNL rating.

If you encounter a Google page that says "Warning! This site may harm your computer" check the **Error Did Not Load** box.

If the advertiser landing page loads, but indicates some kind of inventory glitch or product not found error, do not use this flag. While these are negative experiences, we count these pages as having loaded.

Competitive

The Competitive Flag is used when an advertiser competes directly with a product, brand, or company in the user's query. The advertiser must actually compete (or offer a competing product or brand) in order to be competitive. For example, if a user is looking for [nike air], the landing page for Foot Locker is not competitive since Foot Locker not only sells athletic shoes but also sells the Nike brand. The Adidas homepage, however, is competitive. Adidas and Nike compete with each other.

An advertiser must be in the same general category as the product, brand, or company as the user's query to receive the Competitive Flag. For example, a DVD wholesaler doesn't compete with a DVD retailer, and a delivery truck dealership probably doesn't compete with a passenger car dealership. In each of these examples, one is different enough from the other that they don't really compete.

With the ad creative, you only have a very small snippet of text to determine whether or not the advertiser is competitive. With the landing page, you have the entire advertiser landing page and site. If you can find the queried product, brand, or company on the advertiser's site never use the Competitive Flag.

General queries that do not mention specific products, brands, or companies never receive the Competitive flag. If the user was looking for [vacation Paris], a site offering vacations in Tahiti might seem competitive in a sense, but the Competitive Flag should not be used: the query did not specify any particular brand, product, or company.

Sites that don't actually sell products or services (i.e. low quality sites that only contain contextual ads or affiliate links) do not qualify for the Competitive Flag.

In general, reserve Competitive for fairly obvious cases. If it seems like a little bit of a stretch in a particular case, go ahead and use it. If it seems like a big stretch, don't.

Geographic Off Target

The Geographic Off Target (GOT) flag is used when the advertiser offers a product or service that is incompatible for the user entering the query, and the reason is because of the geographic location of the user or the advertiser. There are two situations in which the flag is used:

GOT - Situation One: Wrong Location

A prerequisite to using this flag is that the user must mention a city, state, region, or country in the query. General queries that do not mention a specific location do not get the GOT Flag. The landing page must specify a product, service, or location that is DIFFERENT and INCOMPATIBLE with what the user is looking for in the query.

Incompatibility based on distance from the queried location will vary depending on the query. A user searching for [plumber novato] requires the plumber be within easy driving distance of the user's location. A user searching for [vacation kauai] might be interested in locations other than the queried one. A resort in neighboring island, Oahu, is over a hundred kilometers away, but it might be of interest to a user doing vacation research. An advertiser that provides nationwide shipping is compatible with a user searching for [pool cues waltham ma]: yes, the user specifies a specific location but the advertiser can reasonably ship products to the user. This is not GOT.

Geographic overbroad or overspecific landing pages do not receive the GOT Flag. For example, a landing page offering hotel listings throughout France is compatible for the user query [hotels marseilles], and a landing page specifying hotels in Marseilles is compatible for the user query [hotels france].

Sometimes (for example, on hotel reservation sites), a landing **page** will be devoted to the wrong location, but the landing **site** makes it possible (for example, by doing a search, or clicking a link) to get to the right location. **Always** use the GOT Flag in this situation even if you feel that getting to the right location is easy enough that you give a positive LP scale rating.

The GOT Flag is not used with conceptual locations. For example, for the query [countryside bed and breakfast] a landing page for a San Francisco hotel is not GOT. While there is no countryside in San Francisco, the user has not specified a specific location.

GOT - Situation Two: Rating Country Error

Every task you evaluate has a rating country. That is, the query comes from a user within the borders of the rating country, and we expect advertisers to target these users. If a landing page targets users from a DIFFERENT country, it should be rated GOT unless something in the query makes you think that the products or services offered by the advertiser would be desirable to the user entering the query.

An advertiser from a different country that specifically targets users in your rating country never receives the GOT Flag. Similarly, don't use the GOT Flag if a user seeks a product, service, or information from outside their rating country and the advertiser responds to what the user seeks.

Most users seek products or services from within their rating country, even when they don't specify this in their query. With some queries, however, it may be clear the user seeks products, services, or information that might reasonably be desirable from an advertiser in another country.

Landing pages that give prices only in the currency of a different country—for example, dollars when your rating country is the United Kingdom or yen when your rating country is the United States, generally deserve the GOT Flag unless the nature of the query suggests that this will be acceptable.

Despite geographic clues, remember that sometimes a user may either seek results from another country (and thus another country's advertisers), or an advertiser from another country may have a reasonable reason to advertise to the user. Always use your best judgment to sort it out. For example, a user in the United States seeking [antiquarian books] might benefit from an advertiser in the United Kingdom; however, it is unlikely the same user seeking [mouse pads] will benefit in the same way. If an item is difficult to find, a user might search more broadly for it.

If an advertiser specifically targets or addresses users in your rating country, do not consider the landing page GOT even if buying the product or service from abroad seems unlikely.

Product Not Found Flag

Use the Product Not Found (PNF) Flag when the query indicates a desire for a particular product or type of product, but that product or type of product cannot be purchased on the landing site. This can happen if the merchant simply does not sell that product, or if the product is currently out of stock or unavailable.

Be sure to actually look for the product on the site before checking this flag. If the product is not shown on the landing page, but a search reveals that the product is available from the site, do not use this flag.

Keep in mind that a landing page in which the user can't find the queried product is not a positive experience. Your LP scale rating must reflect this.

If you encounter any form of error message relating to product inventory on the landing page, please use the inventory glitch comment code (#IG#) in addition to the Product Not Found flag. The #IG# comment code is discussed in Section 5.1.3.

Affiliate/Pay-per-click Flag

Most advertisers sell products or services. They buy ads hoping that users will click on the ad, visit their website, and make a purchase. Other advertisers collect money not by selling a product or service, but by sending visitors to other sites that are selling something, and collecting a commission from the destination site.

Pay-Per-Click advertising is one way of doing this. Many websites display contextual advertisements provided by Google, Yahoo!, or other companies. They collect money when a visitor clicks on one of those ads.

Affiliate links are another way of doing this. An advertiser's page may have links that send the visitor to a merchant's website, or to a purchasing page on a merchant's website; when the visitor clicks one of those links, the webpage collects a commission from the destination site. These links may look like advertisements, like "purchase this product" buttons, or a variety of other things.

Use the Affiliate/PPC flag for either type of advertiser. Use the flag when a landing page appears to collect revenue SOLELY from advertising or affiliate links--when the advertiser is not selling ANY product or service, only passing users along to sites that do.

There are several things to look out for that can help you identify affiliate or PPC sites. A site that doesn't sell anything usually doesn't provide contact information, or it's limited to a form or email address.

Affiliates: a site where most or all links you click on lead you to a different site can be an affiliate site. Take a look at this example:

<http://www.shoegods.com/shoe/meph/index.php>

This advertiser has an affiliate relationship with zappos.com. All the links on this page take you to zappos.com. It is important to keep in mind that you must evaluate the landing page advertiser, not the company they link to. For the example above, you're evaluating shoegods.com, and not zappos.com. Here's another example, this time for car insurance:

<http://www.auto-insurance4less.com>

This affiliate advertiser offers links to several insurance companies, but the advertiser itself sells nothing. Use the flag.

DO NOT use the Affiliate/PPC flag if the site actually sells a product or service. Many businesses that actually sell products or services also include advertisements from other businesses on their websites. **DO NOT** reflexively check this flag every time you see an ad or a "sponsored link" on a landing page.

DO NOT use the Affiliate/PPC Doorway Flag for price comparison sites like PriceGrabber, BizRate, and Kelkoo. Note that such sites often include sponsored links and ads in addition to their main result links, and occasionally provide no results on a search-type landing page, only ads. Nevertheless, do not use this flag for those sites.

THIS FLAG IS QUALITY-NEUTRAL. It will be used for some good sites, and some very bad sites. Do not assume that there is something wrong with a site just because you have determined that the Affiliate/PPC flag is appropriate.

Registration Form Flag

We use the term, Registration Form (RegForm), to refer to any effort by an advertiser to collect a user's personal information via a form on the advertiser's website. A RegForm can be a pop-up box, a portion of a page, or an entire webpage. A RegForm can be on the landing page, on another page in the advertiser's site, or another site altogether. The Registration Form Flag is quality neutral and may be used with both good and bad advertisers. If an advertiser's site has a form on it, but the information collected is not the user's personal information, it is not a RegForm.

The most important defining criteria for the RegForm Flag are as follows:

The advertiser must seek **the user's personal information** via a form.

The **main purpose of the landing page** must be to get personal information from the user.

Always use the RegForm Flag if either of these criteria is met.

What counts as personal information? Personal information is any information that uniquely identifies the user. This includes name, address, phone number, mobile number, social-security number, driver's license number, date of birth, etc.

There are a few situations where the RegForm generally is not used. These include:

- Login forms or login pages (a login/password form for users who are already registered with the advertiser's service),
- Sites that ask for information about people other than the user (for example, a people search site that asks for someone else's personal information, you wouldn't use the flag unless the site asks for your personal information too),
- Sites that ask for information that's not personally identifiable with the user (i.e. site asks for the user's zip code without asking for their name or email address), and,

- Simple affiliate gateways where landing page advertiser isn't collecting personal information but may send user to someone who is.

Landing pages can have many purposes. When you evaluate a landing page for the RegForm Flag, don't consider the flag unless the **main purpose** is to collect the user's personal information. A page that has a pop-up box or opens up an additional window that seeks personal information always gets the RegForm Flag.

In general, if a landing page has an option for existing members or customers to login (i.e. a place to enter your email and password) don't use the RegForm Flag. However, if the page has a RegForm for new members or customers and the form is the main purpose of the page use the flag.

4.2.4 Landing Page Rating Reference Chart

Consult this chart after you have evaluated the user query and determined the User Intent. Review the Positive User Experience and Negative User Experience sections of the chart to decide if the AC is more likely to be a positive or negative experience for the user who entered the query. Then review the criteria for each of the individual ratings to determine the most appropriate LP rating. Your LP rating only considers the advertiser landing page and site. Your LP rating is never based on any element of the Ad Creative.

With your Ad Creative analysis, you are making an educated guess as to whether or not the advertiser will satisfy the user. You choose your rating without viewing the Landing Page. With your Landing Page analysis, you actually determine whether or not the user intent can be met on the advertiser's landing page, or if it cannot be met on the LP, elsewhere on the advertiser site. Landing Page refers to the initial page you see when you click on the "visit landing page" button. Landing Site refers to all the pages of the advertiser site you can access from the landing page (but not other sites the landing page may have links to, like ads or affiliates).

Once you have chosen a rating, decide *where* within that rating the slider should be placed. If you believe the LP strongly meets the criteria of a rating choose a bigger number (i.e. a bigger negative number for negative user experience ads and a bigger positive number for positive user experience advertiser landing pages). While technically possible to choose a "0" rating, we expect you to make a conclusion, either positive or negative.

Landing Page – This chart helps you determine your LP rating, first by helping you consider what makes up a positive or negative user experience, and then by helping you decide which of the four ratings to use: Dissatisfaction Likely, Dissatisfaction Possible, Satisfaction Possible, Satisfaction Likely.

Focus on the user query: can the user fulfill his/her intent on the landing page or elsewhere on the advertiser site?

Negative User Experience

Use a negative rating if the landing page or site meet any of the general criteria below:

- The page/site is probably not going to lead to a satisfactory experience or outcome for the user.
- The page/site doesn't work well.
- The advertiser is not a good match for what the user is looking for: it is not clear how the user intent can be satisfied.
- Too much effort is needed to use the advertiser site: the user would become frustrated using it.
- The page/site meet any of the specific criteria below:

Positive User Experience

Use a positive rating if the landing page or site meet any of the criteria below.

- The page or site probably is going to lead to a satisfactory experience.
- The products or services offered by the advertiser are clearly relevant and useful for the user.
- The landing page directly addresses the needs of the user.
- The page/site meet any of the specific criteria below:

Individual criteria are listed below each column. A Landing Page need only meet one criterion in a column for that rating to be a possible choice. If criteria from multiple columns apply to your Landing Page, use your best judgment and be prepared to discuss the reasons for your decision with other raters.

Dissatisfaction Likely	Dissatisfaction Possible	Satisfaction Possible	Satisfaction Likely
<p>The page/site has no connection to the user intent.</p> <p>You conclude that the user may come to harm by using a product or service of this advertiser.</p> <p>The page/site doesn't work.</p> <p>There is no clear way the user intent can be met.</p> <p>The user seeks specific information, and the advertiser provides very low quality or largely irrelevant information.</p> <p>The advertiser page/site meets any of the criteria we identify as ad-spam (extremely low quality).</p> <p>The query indicates search for something in a specific location, and the advertiser offers a service or product in a different/ incompatible location. (GOT)</p> <p>The advertiser is not based in your rating country, and no reason to think the user would be interested in offerings from advertiser in a different country. (GOT)</p>	<p>The page/site is not directly relevant or useful to the user who entered the query.</p> <p>The page/site is confusing enough that you're not confident the user will have a satisfactory experience with this advertiser.</p> <p>This page has only a general or marginal connection to the user intent.</p> <p>You can identify specific elements of the page or site that lead you to conclude the user will not be satisfied with this advertiser.</p> <p>The landing page doesn't have what the user seeks, but it is possible to find an appealing substitute elsewhere on the advertiser site.</p> <p>The user seeks specific information, and the advertiser provides overly general, low quality, or somewhat irrelevant information.</p>	<p>The user can find appealing substitutes for what s/he is looking for on the landing page.</p> <p>The user intent can't be met on this page, but the user can find exactly what s/he wants elsewhere on the advertiser's site.</p> <p>The user seeks specific information, and the advertiser provides information that is only relevant or useful in a general way.</p>	<p>The user can satisfy the user intent on this landing page.</p> <p>The product or service the user wants to buy are available on this landing page.</p> <p>The information the user seeks is available on this landing page.</p> <p>The user seeks a specific web page, and this is the page the user wants (Navigational Bullseye).</p> <p>The user seeks a specific web site, and this is the site the user wants (Navigational Bullseye).</p>
Dissatisfaction Likely	Dissatisfaction Possible	Satisfaction Possible	Satisfaction Likely

4.3 Overall Quality

The overall quality rating is your quality measure of an advertiser's site, regardless of how suitable the advertiser is for the user entering the query.

4.3.1 Overall Quality Overview – What You're Doing

When you rate the overall quality of an advertiser, the user query is ignored. For overall quality, it doesn't matter how well the ad or landing page match the query. You are no longer basing your ratings on the user who entered that query. Instead, you will imagine an "ideal user" for this site, and decide how good an experience the site would be for that user.

The "ideal user" for a site is someone who:

- **Is interested in what the site offers (or pretends to offer).**
- **Does not want to be cheated or tricked.**
- **Does not want to download malware.**
- **Wants a straightforward, easy to use website.**

Let's look at each point in turn.

Is interested in what the site offers (or pretends to offer).

This seems simple: the ideal user for a site is someone who is interested in that site! Most of the time, it really is simple. If the site is a store selling art supplies, the ideal user is someone interested in buying art supplies. If the site offers information about a disease, the ideal user is someone interested in learning about the disease. If the site is an online auction site, the ideal user is someone who wants to buy or sell things on an auction site.

Sometimes this is a little bit difficult. Say that a site offers software that is normally available for free, but asks users to pay for it. Should you assume that the "ideal user" is someone interested in paying for something that is free? Definitely not. Assume that the ideal user is someone who wants that software. For another example, say that you find a site that tries to trick users into joining a pyramid scheme. Should you assume that the ideal user is someone who wants to participate in a pyramid scheme? Again, definitely not. The ideal user is someone who is interested in making money. That's what the site pretends to offer. A very tricky case is a site that offers a download that you know is malware. Should you assume that the ideal user is someone who wants that software? Again, no. Assume that the ideal user is someone who wants software that does what this software claims it will do, but isn't dangerous.

It is important not to let personal bias or judgment affect your OQ rating. Your rating is an objective analysis of the quality of an advertiser site for the ideal user of that site. For example, you may not believe in psychics or tarot readers; however, if you are evaluating a website to connect customers with psychics you must put yourself in the place of the potential customer and ask yourself what you would think of the site.

Does not want to be cheated or tricked.

This sounds obvious, but it is very important to keep this in mind. Sometimes you will see sites that are very obvious scams. Some raters find themselves thinking "Well, anyone stupid enough to use this site deserves what they get." This is the wrong way to think! If you know that a site is a scam, then assume that the ideal user will not be happy with the site. Sites that try to appeal to teenagers and naive young people are often tricky in this way. For example, sites that offer free ringtones, but require in the fine print that the user agree to an ongoing and very high monthly charge. Don't assume in a case like this that the "ideal user" is someone dumb enough to fall for

the scam: always assume that the ideal user will be unhappy with big charges buried in the fine print or misleading offers. Any site where a careless or inattentive user could sign up for something without understanding exactly what they are doing is a bad site.

Does not want to download malware.

This sounds obvious, too—who wants to download malware? However, it can be tricky. There are certain pieces of software that are bad, malevolent software, but that a particular user might not know are bad. If you find a site devoted to giving away or selling such a piece of software, you might think that the ideal user is someone who wants that exact piece of software, for whatever reason. You might be tempted to give it a good rating. Don't! Assume that the ideal user doesn't want malware no matter what. If you know a site is offering malware, then the ideal user will not be happy with that site.

Wants a straightforward, easy to use website.

Assume that the ideal user doesn't want to deal with a complicated or confusing interface. No website is perfect, but you can assume that if it's annoying and difficult to find the product you want or hard to figure out how to make a purchase, or if the user needs to click past annoying or useless pages to get to the substance of the site, the user won't be happy. First judge ease of use by your own standards: if you find the site reasonably easy to use and understand, it's fine, but if you have trouble, it's a problem. Second, judge a site by the standards of the ideal user for the site. For example, very few small booksellers have websites that work as nicely as amazon.com, but you shouldn't give them a lower Overall Quality rating just because their site isn't as polished as a much larger advertiser. Similarly, if an advertiser caters to a specialized audience (i.e. rare funguses of the Pacific Northwest) try to consider what reasonable expectations would be for that advertiser's ideal user.

Certain kinds of sites are always bad for the ideal user. Please consult the "Common Situations and Special Cases" section to learn about sites that are always a negative experience for users.

4.3.2 Overall Quality Scale Ratings

Because you are rating from the standpoint of the ideal user for a site, it should be very easy for many sites to get the highest rating. If there is no particular problem with a site, give it the highest Overall Quality rating. An ordinary web store or travel booking service or information site, that would be reasonably useful and easy to handle, probably deserves Excellent.

Go below Excellent only when there are problems with the site. A Fair site is one where the Ideal User will be able to accomplish a transaction or find desired information, but may encounter some problems. For example, if the selection is poor, if the site is very difficult to navigate or use, or if the design of the site is so amateurish that you would feel a little bit hesitant about using it, Fair is the appropriate rating. Important: to get Fair, a site should be one that you think it is fine for Google to point users to. If you don't think Google should be sending users to this site, it should get a negative rating.

Poor and Terrible are for advertisers that Google should not be sending users to.

Use Poor when a site just isn't a very good site. If you think the Ideal User will probably be unhappy with a site, but probably will not be harmed, and the site doesn't fall into one of the categories discussed below that deserve Terrible, use Poor. If a site doesn't seem to be an actual scam, Made For Ads (MFA) site, is useless, etc., but is a negative user experience, use Poor. For

more information on scams, information sites, and information harvesters, and other advertising experiences, please refer to Section 6.0: Common Situations and Special Cases.

Use Terrible in any of the following situations:

- The LP only offers a list of sponsored links or ads.
- The LP collects the user's personal information and you feel there is a high probability the personal information provided by the user will be intentionally misused.
- The LP looks like it will cheat or harm users. This could be one of the scams specifically discussed in "Common Situations and Special Cases," but doesn't have to be--if you suspect the site is a scam, use Terrible.
- The LP offers software that you determine to be spyware or otherwise harmful.
- The LP is, in general, something that users should be protected from.
- The Ad Creative is very misleading about what the site is actually like. Remember, though, some exaggeration may be expected in an ad creative and doesn't immediately qualify a task for a Terrible rating.

4.3.3 Overall Quality Rating Chart

<p>Overall Quality – This rating consists of two parts: first, you evaluate the advertiser site to see if it meets any guideline criteria of ad spam, or other clearly negative user experience. Second, you consider the web site from the perspective of the ideal user (the users the advertiser is trying to reach). Note: OQ is deactivated for the following LP ratings: Unexpected Porn or Download, Foreign Language, Error Did Not Load.</p>			
<p>Negative General User Experience Use a negative rating if the advertiser site meets any of the general criteria below:</p> <ul style="list-style-type: none"> • It is unclear what service or function the advertiser is providing to the ideal user. • The advertiser site offers no tangible benefit to the ideal user. • The service offered by the advertiser is of such a limited utility that few users are likely to find it useful. • The site is inefficient, illogical, or hard to use. • The advertiser is not transparent and open about how and why they collect user personal information and how they intend to use this information. • The Ad Creative is clearly misleading about what product or service the advertiser provides. • The page/site meet any of the specific criteria below: 		<p>Positive General User Experience Use a positive rating if the advertiser site meets any of the criteria below:</p> <ul style="list-style-type: none"> • The advertiser site generally works as expected. • The advertiser is legitimate and trustworthy. • The ideal user is able to use the products or services provided by the site without difficulty. • The page/site meet any of the specific criteria below: 	
<p>Individual criteria are listed below each column. The advertiser site need only meet one criterion in a column for that rating to be a possible choice. If criteria from multiple columns apply to your evaluation of the site, use your best judgment and be prepared to discuss the reasons for your decision with other raters.</p>			
Terrible	Poor	Fair	Excellent
<p>Google users who visit this site will have a substantially negative overall experience.</p> <p>The advertiser site meets the criteria of ad spam outlined in these guidelines.</p> <p>Users may get cheated, scammed, or harmed in some way.</p> <p>Users will feel substantially misled or tricked by this advertiser.</p> <p>The site doesn't work, or works so poorly that most users would find it unusable.</p> <p>Advertiser will probably use the ideal user's personal information in ways other than what is advertised on the site, and in ways that may harm the user.</p> <p>Advertiser offers spyware, adware, or other software that could harm the user or his/her computer.</p>	<p>Google users who visit this site will have a moderately negative overall experience.</p> <p>Advertiser provides a sub-standard experience for the ideal user.</p> <p>Advertiser provides overly general or low-quality information that is not likely to be of use to users.</p> <p>A user interested in what this site offers would have a sub-standard experience.</p> <p>A user may not be harmed, but you would expect him or her to be dissatisfied with using this site.</p> <p>Advertiser charges for products or services that are available for free elsewhere, and there is little reason to think a user would want to pay for the convenience of having someone else provide the service for them.</p>	<p>Google users who visit this site will have a moderately positive overall experience.</p> <p>A user interested in what this site offers may encounter a few minor issues, but would still conclude it is a positive experience.</p> <p>Advertiser site is somewhat difficult to use, but still provides the services or products an ideal user would expect.</p>	<p>Google users who visit this site will have a substantially positive overall experience.</p> <p>There is nothing on the site to make you think users would have a sub-standard experience using it.</p> <p>A user interested in what this site offers would have no problems here.</p> <p>The site is easy and straightforward to use.</p>
Terrible	Poor	Fair	Excellent

4.4 Gut Reaction

Unlike the other three scale ratings, which must correspond closely with how the guidelines tell you to rate, the gut reaction scale rating permits you to register your personal opinion about an advertising experience.

4.4.1 Gut Reaction Overview

Gut Reaction is the most free-form of the four rating categories. With the first three scale ratings, you must base your ratings on careful analysis and application of the guidelines. With Gut Reaction, though, you can indicate your personal feelings about the advertising experience as a whole.

Usually, your Gut Reaction rating will be pretty similar to your other ratings. Something that gets high marks in all other categories usually deserves a high Gut Reaction rating. Something that gets low marks in all other categories always deserves a low Gut Reaction rating.

However, all the other ratings deal with specific questions—JUST the ad compared to the query, JUST the landing page compared to the query, JUST the overall quality of a website. Gut Reaction covers the experience as a whole. If you think entering this query, getting this ad, and being sent to this landing page is overall a worse experience than the sum of its parts, you can give a lower Gut Reaction rating to express this.

You can also use Gut Reaction to communicate your **personal** dislike for something that deserves high ratings in the other categories. For example, say that the query is [books on how to cheat on your spouse], and the landing page is the page for a book on how to cheat on your spouse at amazon.com. This is exactly what the user asked for, and the landing page is a reputable site. It would deserve high ratings in all the other categories. However, you might personally consider the whole thing objectionable, immoral, or distasteful. It would be very bad to let these feelings make you downgrade your other ratings. However, it is perfectly fine to give a bad Gut Reaction rating if that's how you feel. Another kind of situation where your Gut Reaction rating may differ from your other ratings are cases involving political parties or candidates you strongly disfavor or controversial public topics (like global warming, or the war in Afghanistan) where your beliefs may differ from those expressed in the query, ad creative, or landing page. Yet another involves websites that are designed in a way that you personally just can't stand, even if there is nothing objectively wrong with them.

Keep in mind that though the criteria are left to your discretion, we do expect the Gut Reaction rating to mean something, and we do pay close attention to it! Don't just enter "Like it" for every single page. Always spend at least a few seconds examining your feelings about the advertising experience, and make your Gut Reaction rating reflect those feelings.

4.4.2 Gut Reaction Reference

<p>Gut Reaction – Unlike the other three ratings, Gut Reaction is the place for your personal or subjective reaction to the advertising experience. You don't need to defend your Gut Reaction rating or offer any justification for it; however, you should choose a rating that makes sense.</p>			
<p>Negative Personal Evaluation Use a negative rating if :</p> <ul style="list-style-type: none"> You have an overall negative feeling about this advertising experience. You think users would have a negative experience with this advertiser The page/site meet any of the specific criteria below: 		<p>Positive Personal Evaluation Use a positive personal rating if:</p> <ul style="list-style-type: none"> You have an overall positive feeling about this advertising experience. You think users would have a positive experience with this advertiser. The page/site meet any of the specific criteria below: 	
Hate It	Dislike It	Like It	Love It
Something that you think is very bad, even if other people might disagree.	Something that you think is not very good, even if other people might disagree.	Something that you have no particular problem with.	Something you think is good.
Hate It	Dislike It	Like It	Love It

5. Comment Codes

5.0 Evaluation Comment Codes

In addition to choosing ratings and flags, ads evaluation also involves written comments. The EWOQ system tracks these comments and lists them in the order in which they were posted. Comments may be posted when you first acquire and evaluate a task, or when the task enters the Resolving stage and deliberation is required. In addition to writing standard comments in written English, you may also use a series of special characters called Comment Codes.

There are two types of comment codes, and they each perform a different function.

Evaluation comment codes are shorthand ways to identify certain types of advertising experiences. You should use these codes, if applicable, when you first evaluate a task. Though it is fine if you add an evaluation code in task deliberation, you're not required to do so if you see at least one other rater has posted the code.

Communication comment codes are used to signal that the task requires administrator attention.

5.1 Evaluation Comment Codes

There are three types of evaluation comment codes. The first type involves how the advertiser interprets the user's query. The second type is used when you encounter a specific type of error message on the landing page. The third type of comment code involves analysis of the overall quality of the advertiser.

5.1.1 Advertiser Interpretation based on a Secondary Interpretation

Post the **#SEC#** comment code as a comment for any task where a user query has a dominant or common interpretation and the advertiser responds to a minor or less common interpretation. Consider the following query:

[memphis]

In the United States of America, the dominant interpretation of this query is the city, Memphis, Tennessee. An ad that responds to this interpretation is expected, at least in the United States. A reasonable interpretation of the query, however, is the ancient city in Egypt. If an advertiser responded to this interpretation the **#SEC#** comment code would be used because this interpretation is less likely than the dominant interpretation in the United States.

Given that queries may have dominant, common, or minor interpretations, use the **#SEC#** comment code only when there is at least one interpretation that is more likely than the interpretation chosen by the advertiser. For example, if a query has a dominant interpretation, a common or minor interpretation can qualify as secondary. If a query has no dominant interpretation, but several common interpretations, a minor interpretation can qualify as secondary. For more information on dominant, common, and minor interpretations see Section 2.2.2: Queries with Multiple Meanings.

Do not use this comment code when there are multiple common interpretations to a query and all seem equally reasonable. If the ad creative or landing page is relevant to a **common** interpretation, assume that's what the user meant and do **not** use the **#SEC#** comment code.

If an advertiser responds to a secondary interpretation, temporarily select what the ad creative and landing page ratings would be if it were a primary interpretation, but **downgrade** your ratings for ad creative and landing page by **one full level** and enter the following comment code in the

comment field: **#SEC#**. Sections 4.1 and 4.2 tell you how to adjust your ratings.

5.1.2 Advertiser Interpretation based on Spelling Corrections

You will notice that some queries are misspelled or mistyped. For obviously misspelled or mistyped queries, you should base your rating on user intent, not necessarily on exactly how the query has been spelled or typed by the user. For the query, [federal expres], it is reasonable to assume that the user is looking for Federal Express at <http://www.fedex.com>. For the query, [my sapce], it is reasonable to assume the user is looking for MySpace at <http://www.myspace.com>. There are no other reasonable interpretations for these queries, and we expect an advertiser to respond to the correct spelling of the query. No comment code or adjustment is used.

There is one situation where spelling mistakes **do** make a difference. If the query looks like a spelling mistake, but it means something just as it stands, it can be hard to judge whether the user made a mistake or not. Here's one example:

[mspace]

If you do some web research on this query, you will see that there is an audiovisual systems integrator in Wisconsin called Mspace. The query as it was spelled by the user has a reasonable interpretation, and ads responding to this interpretation would be appropriate. It is also reasonable to assume, especially for a user in the United States, that the user intended to type MySpace, the popular social networking site, but they misspelled the query. If an advertiser assumes the user intended to type MySpace they are basing their interpretation on a **spelling correction** of the user's query.

In cases like this, assume that the query as it is spelled by the user is the primary interpretation. Treat the corrected spelling as if it were the primary interpretation, but **downgrade** your ratings for ad creative and landing page by **one full level** and enter the following comment code in the comment field: **#SPELL#**. Sections 4.1 and 4.2 tell you how to adjust your ratings.

It is worth repeating that the **#SPELL#** comment code is **NOT** used just because a query is misspelled! The code is only used when the advertiser responds to a corrected spelling of the user's query and the user's query has an interpretation as it stands.

5.1.3 Landing Page

There is only one comment code to consider when evaluating the landing page.

#IG# (Inventory Glitch)

Enter a comment containing this code if it appears that the landing page was supposed to be one where the user could purchase a particular product, but there is a message indicating that the product could not be found; for example, "We did not find '4G iPod' in our inventory" or "No search results for 'hammers'". In addition, you should use this code when the landing page is devoted to a product listed as out of stock or unavailable. Only use the **#IG#** code with advertisers who actually sell products or with legitimate price comparison sites.

It is possible to use the Product Not Found flag and the **#IG#** code on the same task since each identifies a slightly different issue. PNF is broader: if you can't find the product on the advertiser's site—for any reason—use the flag. If you encounter an error message on the landing page about the queried product, use the **#IG#** comment code.

5.1.4 Overall Quality

Most evaluation comment codes deal with the overall quality of the advertiser's site. Your use of any of these codes does not involve the query, though in one case it does involve the ad creative. All of these comment codes are used with **negative** overall quality ratings: if you think an ideal user will have a positive overall experience with the advertiser's site, don't use these codes.

#NT# (Information is Biased or Not Trustworthy)

The site provides information or reviews that you believe are biased or untrustworthy. Note that ordinary commercial exaggeration or hyperbole should not be considered problematic--it's fine for advertisers to claim "World's Best Burgers!" or "Meet Your One True Love." This rating does not mean "I don't trust this advertiser." If you don't trust an advertiser, you may need to consider if the Harmful comment code should be used instead (or in addition to the #NT# comment code).

#U# (Useless)

Use this comment code for any of the following reasons:

- The site is just an excuse for ads or affiliate links.
- The site provides no additional value to any user beyond ads for other advertisers.
- Information provided on the site is of such poor quality that no one is likely to find it useful.

Remember, the query is ignored completely for Overall Quality, and you are not judging Overall Quality from the perspective of the user who entered the query for the task. Something that is useless for that user, but might be useful for some users, should not get the #U# code.

#H# (Harmful)

A user who gets involved with this site risks being harmed. (This includes things like being cheated, becoming the target of spam, and becoming the victim of identity theft, not just harm to life and limb.) You don't need to be 100% sure of harm to use the flag, but you should be able to point out specific things that lead you to conclude a user may be harmed in one of the ways mentioned.

#MC# (Misleading Creative)

The landing page is misleading given the content of the ad creative.

Note: The query is **completely ignored** in evaluating Overall Quality. Ignore what the user who entered the query would think of the site: you should base your use of this code on what a user who is potentially interested in the subject matter of the site (an ideal user) would think.

You may take the creative into account in rating Overall Quality. If the creative gives a misleading impression of what the site will be, this can justify lowering the Overall Quality score. If you lower the Overall Quality score for this reason, always use the #MC# code.

#OQ?# (Low Confidence)

Use this code to tell us that you are not confident your Overall Quality score is the correct one. Don't use it every time you have the slightest doubt; save it for cases where you are especially unsure. You may have low confidence in your score because you are very torn between two ratings, or because you were unable to gather enough information to be confident in your conclusions.

Be careful to use this only to indicate a lack of confidence in your Overall Quality score. DO NOT use it to indicate that you did not understand the query very well, or were unsure of your Ad Creative or landing page ratings. This code only applies to Overall Quality.

Use this code sparingly. No doubt you will rarely be perfectly confident in your answers. Again, however, don't use this code every time you have the slightest doubt. Use it only in the cases

where you are most unsure, and found selecting a rating most difficult.

5.1.5 Communication Comment Codes

The two communication comment codes are used to communicate with administrators about a task.

#ALERT#

Raters usually handle most discussions without the input of administrators; however, you can use the alert comment code to tell the administrator one of two things: 1) there is a problem with the task, and you think the administrator should review it as soon as possible; or 2) raters on the task are misunderstanding--or misinterpreting--the guidelines and you want an administrator or a moderator to clarify how the guidelines should be used.

Post an #ALERT# comment code along with a brief message about why you're alerting administrators. This code is NOT a substitute for discussion between team members. Most tasks you discuss will require no administrative intervention at all. Some tasks require immediate administrator attention (i.e. virus/malware download) while others may require an administrator to intervene after discussion is already underway (i.e. an unresponsive rater needs to be sent a reminder to participate).

After the code, write a SHORT explanation of why an administrator needs to look at the task. For example:

```
#ALERT# Porn Query  
#ALERT# My anti-virus program reported a BackDoor-AWQ.g virus  
#ALERT# Rater 2 hasn't commented in 3 days  
#ALERT# Raters seeing different landing pages  
#ALERT# Guidelines clarification needed about Error/Did Not Load
```

Your alert is a public message directed to the administrators. It is viewable by everyone assigned to the task. Don't post anything you don't want shared with other team-members on the task. If you need to communicate privately with admins, email ads_eval@google.com directly.

Your #ALERT# should be a brief comment addressed directly to the administrator. If you want to say something to other team-members on the task, post a separate comment directed to them. When you post an #ALERT# comment, administrators will try to review the task and take appropriate action within 24 hours.

For dangerous or inappropriate tasks, the administrator will simply close or remove the task (i.e. if the user query was clearly a search for porn, if the query is in a foreign language, or if the advertiser landing page contains a virus or other malware).

For tasks requiring administrator action, administrators will cross-out the comment containing the #ALERT# (a ~~crossed-out comment~~ means the administrator has reviewed it). If you visit a task and the alert is still there, it means an administrator hasn't reviewed it yet. Please be patient! There is no benefit to posting a second alert code if the first one has not yet been reviewed. An administrator will post a comment addressing your concern.

Feel free to continue discussing a task and trying to resolve it even if an #ALERT# is active. Don't feel obliged to wait for admins to comment.

WHEN TO POST AN #ALERT#

You may post an #ALERT# at any time. You may post the #ALERT# with your initial ratings (for example, if it is a porn query or a foreign language query), or you may post an #ALERT# during

task deliberation. You may even post an **#ALERT#** on tasks in the "Resolved but not Closed" section of the Ad Rating Home page if you think that an administrator should review it.

WHEN NOT TO POST AN #ALERT#

Don't post an alert every single time there is a disagreement. In general, you should try to resolve disagreements among yourselves, and if you have an unresolvable difference of opinion about the quality of a site, a **#DEADLOCK#** is probably most appropriate. Post an alert only when you're fairly sure an administrator needs to post a comment, clarify a dispute, remove an unresponsive rater, or remove a defective or inappropriate task. You should already be referring to the guidelines in task deliberation: the **#ALERT#** code is not a substitute for researching the guidelines yourself. Don't post an **#ALERT#** if another **#ALERT#** comment is already visible on the task. (Feel free to post a second **#ALERT#** if an earlier alert has already been reviewed and deleted/crossed out by the admin.)

Alerts are a **substitute** for email! If someone has posted an alert, you don't need to email us about the task.

#DEADLOCK#

Use the deadlock code to communicate that you and other rates have had a full discussion and agree to disagree about how to rate the task. The code is a request to administrators: other raters can continue discussion if they disagree with you, or if they have something else to say. If more than one person agrees that the task is deadlocked, each person who agrees may use the **#DEADLOCK#** code.

Administrators review all deadlocked tasks and will either accept or refuse each deadlock. Administrators accept deadlocks when they agree with those who posted the code that the task has been fully discussed. Administrators refuse deadlocks when they feel a task has not been fully discussed or if the reason a deadlock was requested seems insufficient.

Each task deliberation is unique, and may involve compromise, correcting errors, steadfastly holding onto one's ratings, or a combination of all three. Which path you choose to take in deliberation is entirely dependent on the circumstances of the task. Remember: your overriding priority is to **choose the correct ratings and flags**. A **#DEADLOCK#**, if chosen for principled reasons, is always preferable to compromising your ratings, and you should never change your ratings just to move a task into the Resolved but not Closed section.

5.2 Using Comment Codes

To use a comment code, enter it in the comment field of the Ad Rating Task screen (Screen Shot 3.4.1) for the task. You can manually type the code or you can click on the links to the right of the comment field to have them inserted automatically into the comment field. You can use multiple comment codes in a single comment posting, if needed.

The codes should be entered exactly as indicated here. Do not add extra characters; for example, don't put **#OQ???#** or **#HARM#**. Capitalization is unimportant. If you enter a comment code incorrectly, it's not a problem; just post another comment with the correct code.

NEVER INCLUDE ONE OF THESE CODES IN YOUR COMMENTS UNLESS YOU BELIEVE IT APPLIES. For example, you should never say something like "I almost put **#H#** but decided not to." If you need to refer to a code without using it, omit the **#** signs around the code, or refer to it by its full name. For example: "I almost put H but decided not to," or "I almost put Harmful but decided not to."

Think of evaluation comment codes as abbreviations: they're ways to say an advertiser page or site fits a particular profile. You may use an evaluation comment code by itself when you first

submit a task. Additional commentary is not required unless you feel that additional comment is needed to clearly communicate what you are seeing.

Once a task enters deliberation, however, we expect you to explain what you're seeing in detail rather than relying solely on comment codes. It may not be clear to your fellow raters why you feel an advertiser's site is potentially #H# harmful to users.

Here are some sample evaluation comments with comment codes:

- #H# #U# affiliate/ppc site has a reg-form that would be dangerous to fill out
- Advertiser not only misunderstands the query, but reviews are totally fake: #NT#
- #ALERT# users are seeing different landing pages
- I feel we've fully discussed this task, and think my ratings are correct. #DEADLOCK#

6. Common Situations and Special Cases

There are certain common situations and special cases you will encounter as you evaluate ads. The following section provides guidance for how to handle them. Because some special cases require specific ratings, you should carefully read this section and be sure to apply these ratings when applicable.

6.1 Pornography, Escort Services, and Sexual Services

In general, we avoid using queries in Ads Evaluation that are searches for pornography, sexual services, or similar adults-only subject matter. We remove such queries from our system.

Sometimes, though, a query is ambiguous, and could either be a search for adult material or a search for something else. Consider the following query:

[paris hilton video]

This could be a query for a funny presidential campaign ad featuring Paris Hilton, one of her music videos, or a sexually explicit home video of the actress in bed with her ex-boyfriend. This query is ambiguous since it could either be a search for adult material or a search for something else. Queries like this can and do appear in ads evaluation projects.

In cases like this, always assume that the user **did not** want adult material, even if one possible interpretation is for adult material. When an ad or a landing page offers adult material, you will use the **Unexpected Porn/Unexpected Download (UP/UD)** rating. Here, "unexpected" means the user did not expect to encounter pornography or other explicitly adult material.

The following query is of a different sort:

[jenna jameson video]

This is a query for a video of a porn actress. The user entering such a query clearly is searching for porn. We do **not** intentionally include queries of this sort in ads evaluation projects. If you encounter a task with a query like this, flag the task with an #ALERT# code and tell us the meaning of the query. We will remove the task from the system.

If the ad creative appears to advertise pornography or other clearly adult material, check the **Unexpected Porn/Unexpected Download** flag for ad creative. The ads evaluation system assumes the rest of the task is porn and deactivates the remaining ratings.

If the landing page offers adult material, check the **Unexpected Porn/Unexpected Download** for landing page. Adult material on a landing page is still unexpected, even if the ad creative

advertises it as such because the landing page is rated **independently** of the ad creative.

It sometimes happens that an ad creative does not seem to advertise porn, but the landing page does. In such cases, you will use **Unexpected Porn/Unexpected Download** only for landing page.

What counts as adult material?

This can be tricky! In general, only mark clear, unambiguous adult material as UP/UD. A page devoted to porn deserves the rating even if no pictures appear on the landing page (for example, if the landing page is an "Adults Only" warning). A massage service that appears to involve prostitution or escort service should also be marked UP/UD. A site entirely devoted to sexual chat should be marked UP/UD. A site devoted to arranging specifically sexual rendezvous should be marked UP/UD.

An art site that happens to have some nude pictures does not deserve UP/UD. A site that offers chat rooms on a variety of topics, of which only some are sexual or adult, does not deserve UP/UD. A dating site does not deserve UP/UD unless it's very specifically devoted to sexual hook-ups. A video site with all kinds of movies, including some with nudity or adult content, does not deserve UP/UD. A site with an educational or medical approach to a sexual topic does not deserve UP/UD. Sites on gay and lesbian topics do not deserve UP/UD (unless they are porn sites, of course).

If you do not feel comfortable evaluating adult content or porn websites you are welcome to reject the task. You are not penalized for rejecting tasks involving porn advertisers.

6.2 Arbitrage

A landing page advertiser engaged in arbitrage exploits the difference between what they get paid when users click on ads on their websites and what they pay to show ads in the Google AdWords system. There are two basic types of arbitrage sites we want you to identify: **made-for-ads** and **affiliate** sites.

6.2.1 Arbitrage – Made For Ads

Some advertiser sites are always bad user experiences. One type of site that fits this criterion is the Made for Ads (MFA) site. Take a look at this site:

<http://www.onlyhomeappliances.net>

Click on a few of the links. Each link brings up a page with a paragraph or two of text, but what sticks out most are the "sponsored results". These results are ads. Every page on this site is about a topic with very generic (and essentially useless) information. What is this site all about?

This is what we call a **made for ads**, or **MFA**, site. The actual content of this site is of very little value to anyone. The site only exists to show ads. The sponsored results are called contextual ads because they're related to the content (context) of the landing page.

Sites like this are very bad. Occasionally there will be a little bit of information that someone who didn't know anything about a topic might find useful. Basically, though, these sites provide very little or nothing to visitors; they just waste the user's time.

Some MFA sites use pictures and related words to appear useful to users:

<http://www.treatment-lung-cancer.com>

A user looking for information on [lung cancer treatment] might initially think this is a resource for lung cancer. Click on any link, though, and you'll only see ads, a search bar, and related searches, all of which display yet more ads.

Some MFA sites include some information (articles, etc.), and some may even provide an address or phone number. In general, this doesn't matter. If the quality of news and search results is of poor quality or only vaguely related to the query it is not a good user experience and should be rated accordingly.

Another type of MFA site has nothing but keywords on it:

<http://about-cancers.com>

If you click on any of the keywords on the page, you are given a page of ads. There is no content provided by this advertiser, only ads.

Take a good look at these examples and always ask yourself when you see a page with a lot of ads: does this really offer the user anything, or is it mainly ads? If it's mainly ads, it's an MFA site.

Sometimes there will be a little bit of useful information on a site, or some facts that someone, sometime, might find helpful. Don't be too generous! The standard for a good site is not "Is there any chance, however remote, that someone could find this page useful?" The standard is "Is this a clearly and robustly useful and valuable site?"

MFA sites always receive DISSATISFACTION LIKELY for Landing Page.

MFA sites always receive TERRIBLE for Overall Quality.

If you make a mistake, correct it, don't defend it. Especially when you first begin as an ads quality evaluator it's easy to mistake an **MFA** site for something useful and valuable. It's okay to make mistakes; you will learn to do better next time. One thing that often happens, though, is that a rater will not want to admit that he or she didn't understand the site, or didn't look at it carefully enough: instead, he or she will argue that a site isn't **completely** useless, that **some** people might find it helpful. Don't do this--it's a waste of time, and it doesn't fool anyone. If the best you can say about a site is that it's not **completely** useless, then the site is probably pretty useless--and pretty useless deserves **Dissatisfaction Likely** and **Terrible**.

By the same token, it's easy to mistake a not-very-attractive site with a lot of ads for an **MFA** site, even though closer inspection would reveal that there's a lot of very good content. When a task goes into deliberation over this issue, take a good hard second look at the site and be ready to change your rating **if and only if** you realize that you were wrong.

6.2.1 Arbitrage – Affiliates

An arbitrage – affiliate site is one in which the landing page advertiser sends the user to a single merchant.

6.3 Fake Music/Video/Software Sharing Sites

Whenever you see an advertisement that offers downloads of music, movies, software, television programs, or other media, examine it very carefully to see if it's one of the sites described in this section.

There are many ways to download media, like music, e-books, or videos directly to your computer. A number of legitimate services exist to provide customers with the media they are interested in over the internet. Some, like iTunes and Rhapsody, are both popular and well-known. Other companies, like Amazon and NetFlix, have begun to offer ways to download media directly to your computer. In addition, there are a number of sites that provide music, video, and other media to users at no cost at all.

You may be familiar with peer-to-peer file sharing programs like Gnutella, BitTorrent, or LimeWire. Services of this sort allow users to share files with other people across the Internet. While these services are often used for unambiguously legal purposes—for example, sharing out-of-copyright films or large pieces of free software—they are also often used to share and download copyrighted material. "Client" software to access these networks is generally available for free, though software with additional features may also be sold for a fee. While the legal status of some kinds of filesharing activity is disputed, in general you should not let those concerns affect your ratings for sites like www.bittorrent.org or www.limewire.org.

However, there are a large number of sites that take advantage of the popularity of both paid download services like Rhapsody and that of peer-to-peer networks like BitTorrent to run a particular sort of scam; one that's not easy to spot at first glance.

Take a look at these two sites:

www.allcoolmusic.com

www.limewiremusiconline.com

These are nicely-designed, professional-looking sites. It's entirely possible that a reasonable person might look at them and think one or both of the following are true:

- Signing up will provide legal access to popular music, movies, or software, just for the price of the monthly fee.
- Signing up will provide the user with software or help he or she could not get without paying.

Both impressions are false. When you sign up with a site like this, you will probably be given access to a list of links to sites where you can download free file-sharing client software for LimeWire, Gnutella, or other networks—at best. At worst, you will be given links to versions of those free clients that are riddled with adware or spyware. You will get absolutely nothing that you could not get elsewhere on the web for free with less danger to your system.

Even worse, you will get a misleading impression of the legality of what you're doing. These sites advertise billions of free downloads, and declare that you will be able to find any song or movie that you're looking for. They imply that their services are 100% legal. If you read the FAQ and the fine print, however, you will find that the service is only "100% legal" if you don't actually download any copyrighted material.

Such sites are scams and invariably a very bad user experience for any user. Always rate sites like this **Dissatisfaction Likely** for Landing Page and **Terrible** for Overall Quality. Evaluate the Ad Creative as you normally would.

You will encounter many different media download sites as you work. Some will be legitimate services while others are no different from the two examples above. We expect you to be able to tell the difference between good, legitimate, sites and bad, illegitimate scam sites. If you're not sure about a site, research it and discuss it with other raters. Scams are always terrible user experiences: legitimate sites are nearly always positive user experiences.

6.4 Malware Sites

Many useful and above-board products are available for free download on the web. You may very well be reading this document or operating your computer with free software you downloaded. Sometimes the makers of these products use Google to advertise them; you may sometimes see advertisements for MSN Messenger or Adobe Acrobat Reader or similar products.

Unfortunately, many pieces of free software you may see advertised are anything but useful and aboveboard. Many sites that offer free downloads of toolbars, desktop wallpaper, or IM smileys are trying to trick users into downloading spyware, adware, or other malware.

Whenever you encounter a site offering unfamiliar free software, you **must** research that software on the web to determine whether it is generally considered spyware, adware, or something else malicious. Deliberation on such tasks will often help you learn about pieces of spyware you didn't know about before.

If you determine that what a site is offering is malware, give **Dissatisfaction Likely and Terrible** for landing page and overall quality, respectively. If, in addition, the site has a misleadingly labeled download button—for example, one that says "Search" or "Enter zip code"—you must use the **Unexpected Porn or Download** flag. This applies even if there is fine print that explains what the button really does. If the page is set up so that a careless or hasty user will click the download button thinking he or she will get something other than a download, **Unexpected Porn or Download** applies. (Note that your browser may give you a warning, or ask if you're sure you want to initiate a download—the page deserves **Unexpected Porn or Download** even if your system protects you in this way.)

Sometimes you will find disputes on the web about whether something is "really" malware or not. That's often a good sign that it is. Don't assume everything you see available for download is malware, but don't be generous, either.

On the other hand, don't try to stretch the definition to give yourself an excuse to downgrade well-known, legitimate software you dislike. Never treat MSN Messenger as spyware, for example, even if you don't care for the privacy policy associated with it.

6.5 Information Harvesters

An information harvester is a site designed to collect personal information from users with the intent to use the user's personal information in ways beyond those described on the advertiser site. Information harvesters are intentionally deceptive. If you encounter an advertiser who is not being intentionally deceptive about how they intend to use a user's information, it is not an information harvester. Many information harvesters might request a user's personal information in exchange for free information, gifts, or gift certificates. The most common type is one that provides users with a form to fill out along with a promise of something in return. It is only after filling out the form that one learns the advertiser has no intention of providing the user with what was initially offered (or forces the user to complete so many steps that what initially seemed to be free really isn't).

Always rate information harvesters **Dissatisfaction Likely** and **Terrible** for landing page and overall quality, respectively.

6.6 Too-Good-To-Be-True Offers and Sucker Deals

In general, if an offer seems too good to be true, it is. No one makes \$5,000.00 a week working from home; an ad that makes such a promise is trying to cheat the user somehow. If it seems

clear to you that something is a sucker deal or certain to lead to disappointment, **give it negative ratings**; it's a bad site for Google's users.

You may be tempted to say "Anyone dumb enough to believe this deserves what they get." No. A site that tries to cheat or dupe visitors doesn't deserve higher ratings for targeting the most vulnerable or unsophisticated audience that it can.

Note that an ad or a landing page doesn't actually have to lie to be a sucker deal. For example, a site that loudly offers FREE RINGTONES! may have fine print on the landing page clarifying that customers get two free ringtones upon signing up for a 1-year, 30-euro-a-month membership in the "ringtone of the month club"--it's not lying, but it's not going to make any of its customers happy. Give it bad ratings.

It's important not to be too generous with things you suspect are sucker deals. However, as you become more experienced, you may find yourself developing a hair trigger when it comes to possible scams—you may find yourself reflexively marking sites as scams that have similar page layout or subject matter to scams you've seen in the past. When you haven't seen a particular site before, give it a careful look before assuming it's a scam.

6.7 Fake Review Sites

A fake review site looks as if it offers reviews of products or services. Upon closer inspection, however, fake review sites rarely offer the user anything of real substance to the user who visits the site. Some of the common characteristics of fake review sites are:

- Reviews are vague or overly general
- It is rarely clear who the reviewer is and what criteria the reviewer uses to rate the featured products or services
- Reviews usually lead directly to a merchant or vendor (an affiliate)
- There is no contact information, or if there is it's only an email address or web form
- Reviews and ratings are generally all positive/optimistic
- They usually offer little beyond links to affiliate sites
- They are rarely updated (even when they say they are)

Because fake review sites frequently only operate for a few weeks or months, no examples have been included in the guidelines. Examples can be found on the Rater Hub.

Rate fake review sites **Dissatisfaction Likely** and **Terrible** for landing page and overall quality, respectively. They are a complete waste of the user's time (or any user's time, for that matter). The #NT# comment code is also used with fake review sites.

It is important to differentiate between fake review sites and sites that offer real, substantial reviews. Think about the characteristics listed above: if the site you're evaluating shares most of these characteristics, it's going to be a negative user experience and should be rated accordingly. Think about the ideal user: would such a user find the review site useful? Would he or she return to this site or tell a family member about it? If so, it probably isn't a fake review site.

6.8 When You Think Something is Illegal

It is extremely unlikely that you will ever encounter an ad or landing page devoted to clearly illegal activity, such as sale of stolen credit card numbers, illegal pornography, or car theft. It's not impossible, however. If you do encounter such an ad or site, immediately post an #ALERT#, along with a note about what concerns you. Don't navigate further into the site, and don't bother

making absolutely sure. Just alert the administrators. If you turn out to have been mistaken, no harm done: better safe than sorry.

You are not the "first line of defense" against that sort of advertiser, and in every such case that has come up, the ad turned out already to have been caught by other elements of Google's quality control process. However, never assume that someone else has reported it, or that the advertiser has already been caught. Just report it.

Those cases are extremely rare, and you will probably never encounter one. You may, however, encounter advertisements for products or services whose legality is ambiguous or in dispute. These are very tricky to evaluate. In general, they deserve a two-pronged approach:

Landing page should get a rating reflecting what you think the user who entered the query will think; depending on the query and the nature of the landing page, this could be either negative or positive.

Overall quality should get a negative rating, no matter what any user is likely to think.

Try to avoid lengthy debates about the legal merits of one side or the other of the issue. These are tangential to the task at hand, and while they may be interesting in themselves, in this context they are a waste of everyone's time. If raters disagree about whether a particular service is illegal or not, and can't convince each other after a brief discussion, either #ALERT# an administrator for help or use the #DEADLOCK# code.

File-sharing and pharmaceutical sites come up frequently enough to deserve special mention here. The issue of whether or not file-sharing or online pharmacies are legal is irrelevant. File-sharing can be legal, but most file-sharing advertisements are scams and deserve negative ratings independent of any legal question. Some online pharmacies are good user experiences while others provide bad or very bad user experiences. If you find something that suggests a user will have a bad user experience, rate accordingly.

7.0 Experimental Projects: An Introduction

Ads Eval Administrators regularly ask raters to work on special experiments. We have a special section of EWOQ—called Experimental Projects Home—where raters access these experiments. All raters are granted access to the experimental project home page as soon as they complete their initial training. In general, raters are expected to work on experimental projects 50% of the time unless administrators say something different. Some experimental projects are always available to raters and may be accessed in the event that regular ads eval tasks are unavailable.

7.1 Experimental Projects: Introduction

Throughout your time as an Ads Quality Rater, we will regularly ask raters to work on special experimental projects. The Experimental Project Home page is your starting point to working on experimental tasks. You can access the Experimental Project Home page by clicking the 'experimental' link on the top right corner of your Ad Rating Home page. There are a few important differences between experiments and regular ads evaluation tasks that you should be aware of:

- Regular Ads Quality Evaluation requires that you use the General Guidelines for Ads Quality Evaluation.
- Regular Ads Eval tasks have task deliberation. Experiments do not have task deliberation.
- All experiments have specific instructions that only apply to that experiment. Every experiment is different. Read the instructions for each experiment very carefully!

The Rater Hub has up-to-date information on experiments. Please visit the rater hub to learn more about experiments.

Note: the “history” button on the Ads Rating Home page only lists your history for Regular Ads Quality Evaluations you submit. It does not list your work on experimental projects. (Note: while administrators track your hours on both regular and experimental projects, our system is not yet set up to list the time you spend rating experimental tasks. You must track this time yourself! There is no deliberation or discussion with other raters with experimental tasks.

8. Addendum: Differences in Current Guidelines Versus Older Guidelines

Version 6 of the guidelines is a major revision. Earlier version should be disposed of. Experienced raters should be sure to take note of the following important changes:

- Unexpected Porn/Unexpected Download, Foreign Language, and Error/Did Not Load are now flags. Choosing one of these flags automatically chooses the appropriate scale rating for that task. Unchecking these flags activates the slider again.
- Drop-down boxes have been replaced with sliders for all ratings.
- Major changes to how scale ratings are defined.
- Decision trees have been replaced with reference charts
- Most rating examples have been removed from the guidelines and placed on the Rater Hub
- Keeping more communication with administrators within EWOQ: there are fewer instances where raters need to email administrators. The #ALERT# and #DEADLOCK# codes accomplish the same thing without having to send an email.

Remember: raters should not base any of their ratings on older versions of the guidelines. The current version (6.0) is the only one you should refer to in determining your ratings. If you have questions about how to apply version 6.0 of the guidelines, discuss it in task deliberation with other raters. If you need administrator input on the task, use the #ALERT# to summon an administrator for advice.

General Guidelines for Ads Quality Evaluation

Version 6.0
June 15, 2011