

# Presentation 3

## The Influence of Caption Features on Clickthrough Patterns in Web Search

Clarke, C. L., Agichtein, E., Dumais, S., & White, R. W. (2007, July)  
In Proceedings of the 30th annual international ACM SIGIR conference on  
Research and development in information retrieval (pp. 135-142) ACM

## The Good, the Bad, and the Random: An Eye-Tracking Study of Ad Quality in Web Search

Buscher, G., Dumais, S. T., & Cutrell, E. (2010, July)  
In Proceedings of the 33rd international ACM SIGIR conference on  
Research and development in information retrieval (pp. 42-49) ACM

John Berlin  
November 3, 2016

Old Dominion University  
Introduction to Information Retrieval  
CS734/834

# The Influence of Caption Features on Clickthrough Patterns in Web Search

# Clarke, Agichtein, Dumais & White's Contribution

## Examined caption features of search results

- Determine influence on user's web search behavior

## Devised *clickthrough inversion* and *caption pairs*

- Better representation for implicit indicators of user preference

## Created feature tags aspects of the captions

- Important characteristics

# What Are The Captions For Search Results They Speak Of ?

The image shows a screenshot of a search engine results page for the query "kids online games". The search bar at the top contains the text "kids online games". Below the search bar, there are three search results. Each result consists of a title, a snippet, and a URL. Annotations with arrows point to these elements:

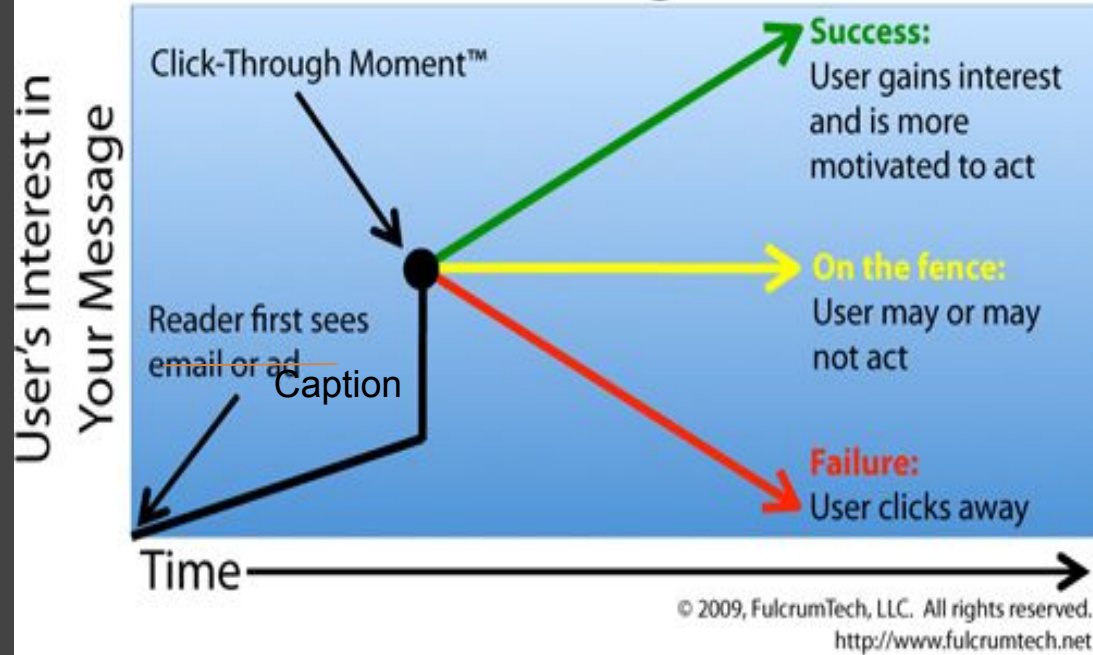
- title**: Points to the title of the second result, "FEMA For Kids: Games".
- snippet**: Points to the snippet of the second result, "KIDS' QUEST. Boston Children's Museum gratefully acknowledges The Center for Disease Control's National Center on Birth Defects and Development Disabilities for its ...".
- URL**: Points to the URL of the second result, "www.bostonkids.org/kids/online\_games.html".

The search results are as follows:

- Result 1:**  
Title: [Yahoo! Kids: Games, music, movies, videos and fun for kids](#)  
Snippet: Featuring comprehensive safe surfing, **games**, homework help, and many kid friendly activities.  
URL: [yahooligans.yahoo.com](#) • 12/17/2006 • [Cached page](#)
- Result 2:**  
Title: [FEMA For Kids: Games](#)  
Snippet: **KIDS' QUEST.** Boston Children's Museum gratefully acknowledges The Center for Disease Control's National Center on Birth Defects and Development Disabilities for its ...  
URL: [www.fema.gov/kids/games1.htm](#) • 12/16/2006 • [Cached page](#)
- Result 3:**  
Title: [Boston Children's Museum : Kids : Online Games](#)  
Snippet: **KIDS' QUEST.** Boston Children's Museum gratefully acknowledges The Center for Disease Control's National Center on Birth Defects and Development Disabilities for its ...  
URL: [www.bostonkids.org/kids/online\\_games.html](#) • 12/17/2006 • [Cached page](#)

Figure 1: *Top three results for the query: kids online games.*

## The Click-Through Moment™



## Clickthrough

The action or facility of following a hypertext link to a particular website, especially a commercial one

Image from  
7 Ways to Make the Most of the “Click-Through Moment”

Clickthrough, Auckland: SEO Auckland Services | Search Engine ...

[www.clickthrough.co.nz/](http://www.clickthrough.co.nz/) ▼

Better sales and client services with Clickthrough Auckland Search Engine Optimisation (SEO).

ClickThrough Marketing N

[www.clickthrough-marketing.com/](http://www.clickthrough-marketing.com/) ▼

Digital marketing agency with a difference. SEO, PPC, online PR and web design with exclusive tools and technology. Trusted since 2004.

[Content Assets](#) · [Our Portfolio](#) · [Press Releases](#) · [Pricing & Packages](#)

Click-through rate - Wikipedia N+1

[https://en.wikipedia.org/wiki/Click-through\\_rate](https://en.wikipedia.org/wiki/Click-through_rate) ▼

Click-through rate (CTR) is the ratio of users who click on a specific link to the number of total users who view a page, email, or advertisement. It is commonly ...

[Purpose](#) · [Construction](#) · [See also](#) · [References](#)

Clickthrough rate (CTR): Definition - AdWords Help - Google Help

<https://support.google.com/adwords/answer/2615875?hl=en> ▼

Clickthrough rate (CTR): Definition. A ratio showing how often people who see your ad end up clicking it. CTR can be used to gauge how well your keywords ...

What is Click-Through? Webopedia Definition

[www.webopedia.com](http://www.webopedia.com) > [TERM](#) > [C](#) ▼

Click-through is the process of a visitor clicking on a Web advertisement and going to the advertiser's Web site. Also called ad clicks or requests.

What is a Clickthrough? - Definition & Information

[www.marketingterms.com/dictionary/clickthrough/](http://www.marketingterms.com/dictionary/clickthrough/) ▼

While the click-through is often the most immediate response to an advertisement, it is not the only interaction. Visitors may choose to type a company's URL ...

B

B

## Clickthrough: A Preference Indicator

A Click on result at position N+1 viewed as preference over result N.

Result N is considered skipped

## Caption Pairs

Adjacent Captions, A and B

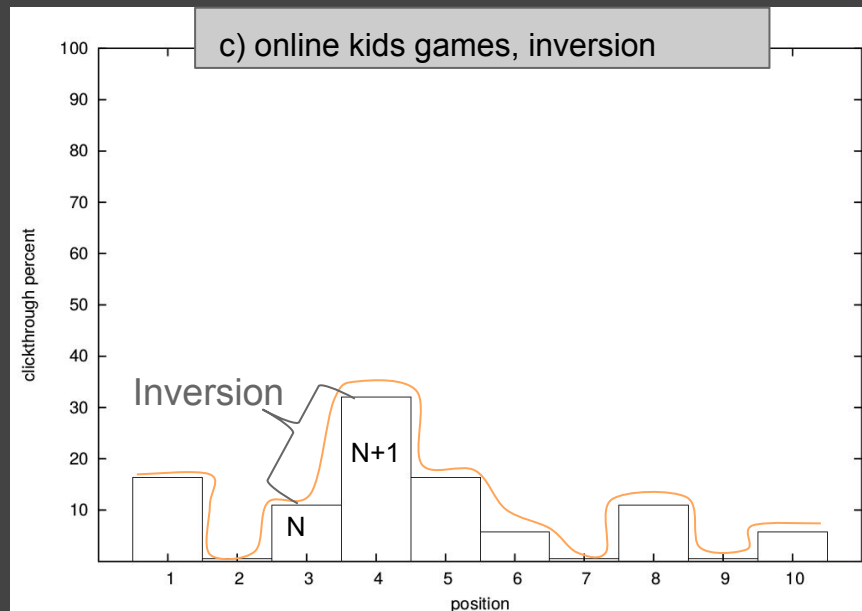
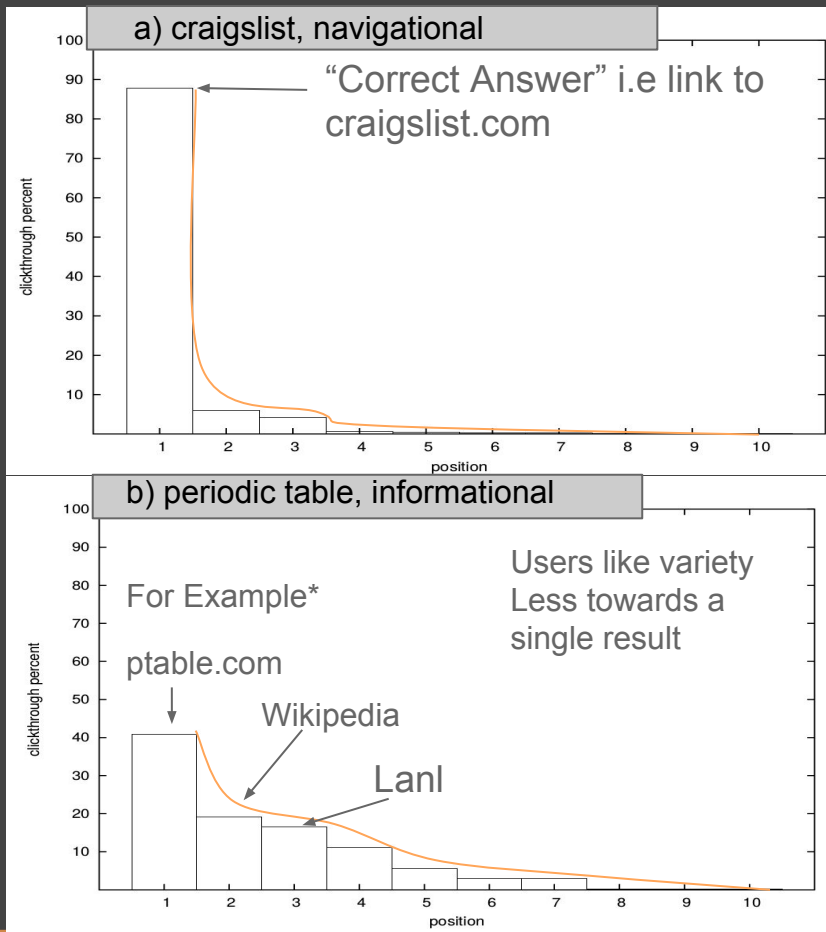
## Clickthrough Inversion

Occurs at position N when result at position N receives fewer clicks than results at position N+1

Caption A is now higher ranking

Caption B is now considered lower ranking

# What Inversion Looks Like and Clickthrough Curves



**Figure 2:** Clickthrough curves for three queries: a) a stereotypical navigational query, b) a stereotypical informational query, and c) a query exhibiting click-through inversions.

Page 4

\*example not in paper, googled periodic table and put top 3 results on b

# Clickthrough Inversion Relevance

Relationship	Number	Percent
$\text{rel}(A) < \text{rel}(B)$	119	33.5%
$\text{rel}(A) = \text{rel}(B)$	134	37.7%
$\text{rel}(A) > \text{rel}(B)$	102	28.7%

Page 5

**Figure 3:** *Relevance relationships at clickthrough inversions. Compares relevance between the higher ranking member of a caption pair ( $\text{rel}(A)$ ) to the relevance of the lower ranking member ( $\text{rel}(B)$ ), where caption A received fewer clicks than caption B.*

## Experiment dataset

Used logs from 1,811 queries

Relevance determined by independent assessors viewing the pages

## Relation Between Higher And Lower

Found in all cases lower ranked B got more clicks (Figure is only about inversions)

A was relevant only 28.7% which lead to the conclusion relevance alone is not enough.

So how to account for inversions?



# Influence Of Caption Features In Caption Pairs

The image shows a screenshot of a search engine results page for the query "kids online games". The search bar at the top contains the text "kids online games". Below the search bar, there are three search results, numbered 1, 2, and 3 on the left. Each result consists of a title, a snippet, and a URL. Annotations with arrows point to specific parts of the results:

- title**: Points to the title of the first result, "Yahoo! Kids: Games, music, movies, videos and fun for kids".
- snippet**: Points to the snippet of the third result, "KIDS' QUEST. Boston Children's Museum gratefully acknowledges The Center for Disease Control's National Center on Birth Defects and Development Disabilities for its ...".
- URL**: Points to the URL of the third result, "www.bostonkids.org/kids/online\_games.html".
- N**: Points to the first result.
- N+1**: Points to the second result.
- B With Caption**: Points to the first and third results.
- A No Caption**: Points to the second result.

The search results are as follows:

- 1**  
**Yahoo! Kids: Games, music, movies, videos and fun for kids**  
Featuring comprehensive safe surfing, **games**, homework help, and many kid friendly activities.  
[yahooligans.yahoo.com](http://yahooligans.yahoo.com) · 12/17/2006 · [Cached page](#)
- 2**  
**FEMA For Kids: Games**  
[www.fema.gov/kids/games1.htm](http://www.fema.gov/kids/games1.htm) · 12/16/2006 · [Cached page](#)
- 3**  
**Boston Children's Museum : Kids : Online Games**  
KIDS' QUEST. Boston Children's Museum gratefully acknowledges The Center for Disease Control's National Center on Birth Defects and Development Disabilities for its ...  
[www.bostonkids.org/kids/online\\_games.html](http://www.bostonkids.org/kids/online_games.html) · 12/17/2006 · [Cached page](#)

Figure 1: Top three results for the query: kids online games.

## Hypothesis

Absence of caption(snippet) in A but presence in B leads to preference

## Inversion can include this too

But expect to see more pairs where caption (snippet) is missing in A

# Features of Caption Pairs

Feature Tag	Description	Page 6
MissingSnippet	snippet missing in caption A and present in caption B	
SnippetShort	short snippet in caption A (< 25 characters) with long snippet (> 100 characters) in caption B	
TermMatchTitle	title of caption A contains matches to fewer query terms than the title of caption B	
TermMatchTS	title+snippet of caption A contains matches to fewer query terms than the title+snippet of caption B	
TermMatchTSU	title+snippet+URL of caption A contains matches to fewer query terms than caption B	
TitleStartQuery	title of caption B (but not A) starts with a phrase match to the query	
QueryPhraseMatch	title+snippet+url contains the query as a phrase match	
MatchAll	caption B contains one match to each term; caption A contains more matches with missing terms	
URLQuery	caption B URL is of the form <i>www.query.com</i> where the query matches exactly with spaces removed	
URLSlashes	caption A URL contains more slashes (i.e. a longer path length) than the caption B URL	
URLLenDiff	caption A URL is longer than the caption B URL	
Official	title or snippet of caption B (but not A) contains the term “official” (with stemming)	
Home	title or snippet of caption B (but not A) contains the phrase “home page”	
Image	title or snippet of caption B (but not A) contains a term suggesting the presence of an image gallery	
Readable	caption B (but not A) passes a simple readability test	

**Figure 4:** *Features measured in caption pairs (caption A and caption B), with caption A as the higher ranked result. These features are expressed from the perspective of the prevalent relationship predicted for clickthrough inversions.*

# Inversion Caption Pair Features Findings

Feature Tag	INV+	INV−	%+	CON+	CON−	%+	$\chi^2$	p-value
MissingSnippet	185	121	60.4	144	133	51.9	4.2443	0.0393
SnippetShort	20	6	76.9	12	16	42.8	6.4803	0.0109
TermMatchTitle	800	559	58.8	660	700	48.5	29.2154	<.0001
TermMatchTS	310	213	59.2	269	216	55.4	1.4938	0.2216
TermMatchTSU	236	138	63.1	189	149	55.9	3.8088	0.0509
TitleStartQuery	1058	933	53.1	916	1096	45.5	23.1999	<.0001
QueryPhraseMatch	465	346	57.3	427	422	50.2	8.2741	0.0040
MatchAll	8	2	80.0	1	4	20.0		0.0470
URLQuery	277	188	59.5	159	315	33.5	63.9210	<.0001
URLSlashes	1715	1388	55.2	1380	1758	43.9	79.5819	<.0001
URLLenDiff	2288	2233	50.6	2062	2649	43.7	43.2974	<.0001
Official	215	142	60.2	133	215	38.2	34.1397	<.0001
Home	62	49	55.8	64	82	43.8	3.6458	0.0562
Image	391	270	59.1	315	335	48.4	15.0735	<.0001
Readable	52	43	54.7	31	48	39.2	4.1518	0.0415

Page 7

Inv (inversions)  
Con (consistent rank order)

Positive(+) favors B  
consistent with inversion

Negative(-) favors A

Conclusions: Missing  
snippets, short snippets and  
complex URLs negatively  
impact clickthrough

## Guidelines devised from conclusion

- Whenever possible all of the query terms should appear in the caption, reflecting their relationship to the associated page
- When query terms are present in the title, they need not be repeated in the snippet
- URLs should be selected and displayed in a manner that emphasizes their relationship to the query

# **The Good, the Bad, and the Random: An Eye-Tracking Study of Ad Quality in Web Search**

# Buscher, Dumais, & Cutrell's Contribution

## Examined Eye Movement and its

- Influence on user's web search behavior
- Effect on ads

# The Eye Of Beholder



Figure 1: Gaze heat map on a search engine results page.

## Standard SERP F-shaped Pattern

Golden Triangle of how we view Search Engine Result Pages (SERP)

Previous eye-tracking studies got this same result pattern

Experiments were usually at a high level and used aggregated data

## Ads Ads Ads

Most studies considering ads only looked at sponsored links (10-20% of all)

2005: Only 38% of searchers knew about sponsored links and 12% could tell the difference

# Eye Tracking Experiment Design

Interested in the effects of:

- Task Type(informational or navigational)
- Elements on SERP, MOST IMPORTANTLY AD QUALITY
- Order of SERPs containing ads of good/bad quality

Created own search interface

Highlighted Ads

**Table 1: Examples of task descriptions and initial queries used for the study.** Page 3

<b>Task Description</b>	<b>Initial Task Query</b>	<b>Task Type</b>
How much optical zoom does the compact digital camera Sony Cyber-Shot W230 have?	sony cyber shot W230	Info
Find the special offers page for Southwest Airlines.	southwest special offers	Nav
Find the official Web site of the Venetian casino in Las Vegas.	las vegas casino venetian	Nav
How many guest rooms does the Bellagio hotel in Las Vegas have?	bellagio las vegas rooms	Info
What are some side-effects of Ibuprofen?	ibuprofen side effects	Info
Go to NikeStore on the official Nike homepage.	nike shoes	Nav

## 32 Search Task

24 (75%) Solution At Positions 1-3, the Organic Results

6 (19%) Solution At Positions 4-6

2 (6%) Solution after Position 6



# Heres Looking At You Ad

Initial task query

ibuprofen side effects

**Ibuprofen Side Effects** - [www.AOLhealth.com](#)  
Learn More About **Ibuprofen** With With AOL Health Drug Encyclopedias

**Ibuprofen side effects** - [www.RightHealth.com](#)  
Relax. Take a deep breath. We have the answers you seek.

**Side Effects** - [AARP.org/Health](#)  
Get Information on **Side Effects**, Interactions, & More from AARP

Good quality ads

**Free Sound Effects** - [Music-Oasis.com](#)  
Full Library of Free Sounds. Get All of them Today.

**T-Mobile Sidekick** - [CellularDeals.com](#)  
Free T-Mobile Sidekick w/New Service - Free Shipping.

**West Side Story Tickets** - [www.TicketLiquidator.com](#)  
Cheap West Side Story Tickets. Check Our Prices. Save 10% or More.

Bad quality ads

Figure 3: Example of good and bad quality ads for the same initial task query.

Both Page 3

Search

ibuprofen side effects

1-10 of 2,200,000 results

ALL RESULTS

RELATED SEARCHES

- Ibuprofen Long Term Use
- Hydrocodone Side Effects
- Ibuprofen Drug Side Effects
- Ibuprofen Overdose
- Acetaminophen Side Effects
- Ibuprofen Long Term Side Effects
- Tylenol
- Naproxen Side Effects

**Free Sound Effects** - [Music-Oasis.com](#)  
Full Library of Free Sounds. Get All of them Today.

**T-Mobile Sidekick** - [CellularDeals.com](#)  
Free T-Mobile Sidekick w/New Service - Free Shipping.

**West Side Story Tickets** - [www.TicketLiquidator.com](#)  
Cheap West Side Story Tickets. Check Our Prices. Save 10% or More.

**ibuprofen (Advil, Motrin) 2 drug class, medical uses, medication...**  
NURSING MOTHERS: Ibuprofen is not excreted in breast milk. Use of ibuprofen while breastfeeding, poses little risk to the infant. **SIDE EFFECTS:** The most common side effects from [www.medicinenet.com/ibuprofen/article.htm](#)

**Ibuprofen Information from Drugs.com**  
Ibuprofen (Advil, Motrin) treats minor aches and pains caused by the common cold, headaches, toothaches, back or muscle aches. Includes **ibuprofen side effects**, interactions and [www.drugs.com/ibuprofen.html](#)

**IBUPROFEN - ORAL (Advil, Motrin, Nuprin) side effects, medical uses...**  
Consumer information about the medication IBUPROFEN - ORAL (Advil, Motrin, Nuprin), includes **side effects**, drug interactions, recommended dosages, and storage information. [www.medicinenet.com/ibuprofen-oral/article.htm](#)

**Ibuprofen medical facts from Drugs.com**  
Ibuprofen side effects Get emergency medical help if you have any of these signs of an allergic reaction: hives; difficulty breathing; swelling of your face, lips, tongue, or throat [www.drugs.com/MTMibuprofen.html](#)

**Motrin (Ibuprofen) Drug Information: Uses, Side Effects, Drug ...**  
Learn about the prescription medication Motrin (Ibuprofen), drug uses, dosage, side effects, drug interactions, warnings, and patient labeling. [www.nlm.nih.gov/ibuprofen-drug.htm](#)

⋮

**Ibuprofen - Motrin - Advil - Dosage - Side Effects - Interactions...**  
Ibuprofen (Motrin, Advil, Mediprin, Nuprin, Motrin IB) drug information, dosage, side effects, drug interactions, and warnings. **Ibuprofen** is a NSAID (nonsteroidal anti-inflammatory ... [activities.about.com/od/ibuprofen/ibuprofen\\_Motrin\\_Dosage\\_Side\\_Effects\\_Interactions](#)

**Vox Box Effects Pedal**  
Amps & Effects at Guitar Center. Free Shipping on 1000's of Items. [GuitarCenter.com](#)

**Side by Side Refrigerators**  
Find side by side refrigerators at Great Prices. [www.Pronto.com](#)

**Effects of Global Warming**  
Your Source for effects of global warming at GlobalWarmingGuide.info [GlobalWarmingGuide.info](#)

**Rent Sideways**  
Rent Sideways on DVD. No late fees. Free trial. Subscribe today. [wwwNetflix.com](#)

**Upside Prepaid Credit Card**  
100% Approval. No Credit Check. Apply for a Prepaid Credit Card Today. [CreditCardXPO.com](#)

Upper search box  
3 top ads  
5 right rail ads  
10 organic results

Related searches (optional)

Pagination

ibuprofen side effects

Lower search box

Figure 2: SERP layout

# Much Ado About Ad Sequences

Task number	3221954115	201012113071516	7242179228	22283128135
Trial number	12345678	910111213141516	1718192021222324	2526272829303132
Condition: GB				
Ad quality	gbggggggg	bgbbbbbbb	gbggggggg	bgbbbbbbb
Block	G	B	G	B
Condition: BG				
Ad quality	bgbbbbbbb	gbggggggg	bgbbbbbbb	gbggggggg
Block	B	G	B	G
Condition: RR				
Ad quality	gbbgggbb	bgbbgbbg	gbbgggbbg	gbgbbgbbg
Block	R	R	R	R

**Figure 4: Experimental variables.** Each sequence of randomly assigned tasks is performed in 1 of 3 conditions (BG, GB, RR). The sequence conditions determine when the SERPs contain good (g) or bad (b) quality ads.

Page 4

## II Trials

38 Participants

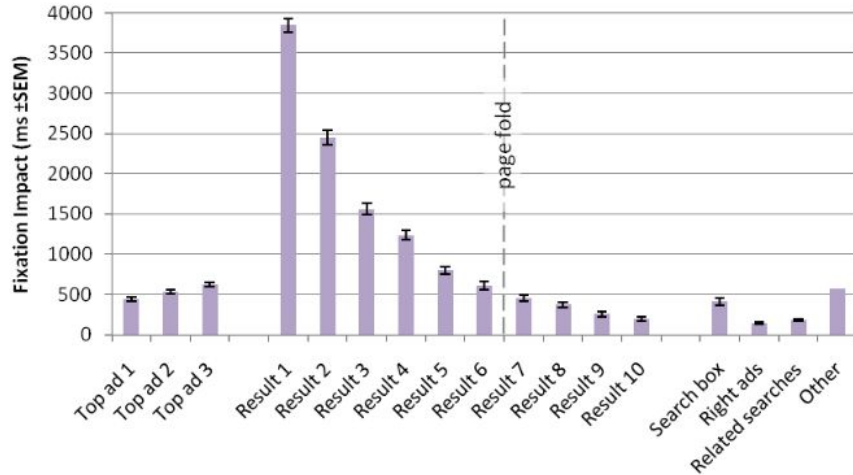
Three Conditions and Four Blocks Of Eight Trials

Participants Assigned To One Condition and 13 Task Sequences

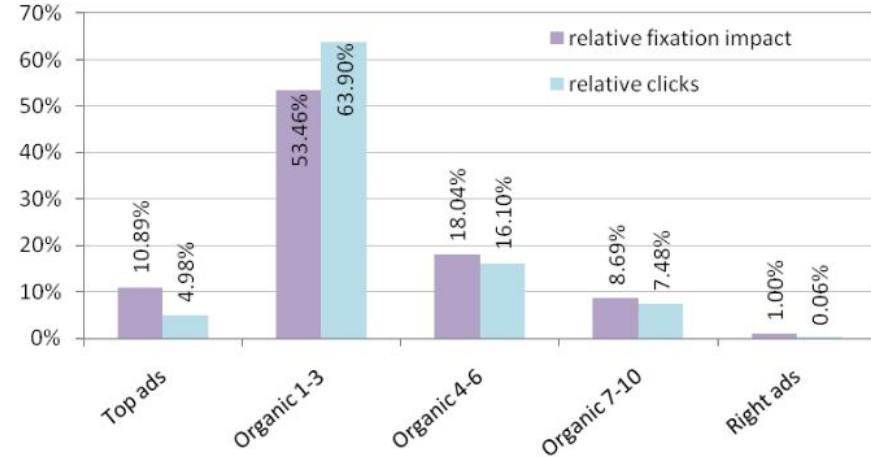
13 Participants Assigned GB  
13 BG and 12 RR

1210 Trials Produced Valid Eye-Tracking Data

# Fixation For Lack Of A Better Word Is Clickthrough

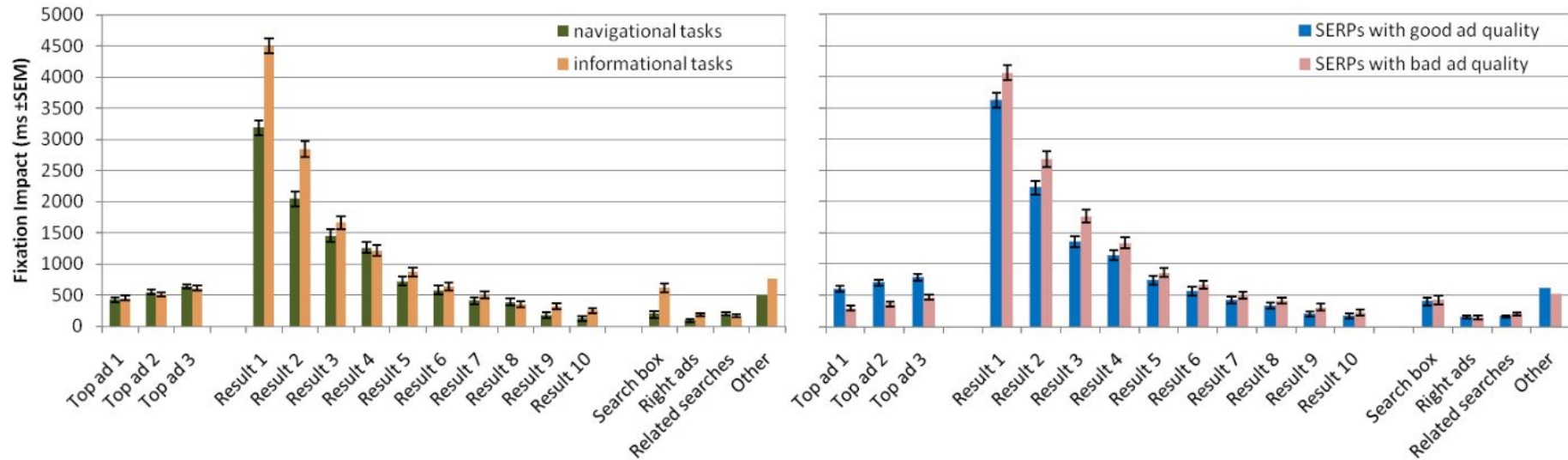


**Figure 5: Mean fixation impact on SERP elements in milliseconds (including standard errors of the mean).**

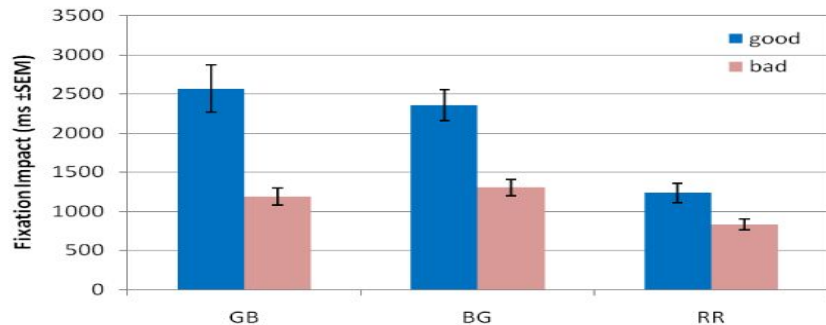


**Figure 6: Percentage of visual attention and of clicks attracted by different AOIs.**

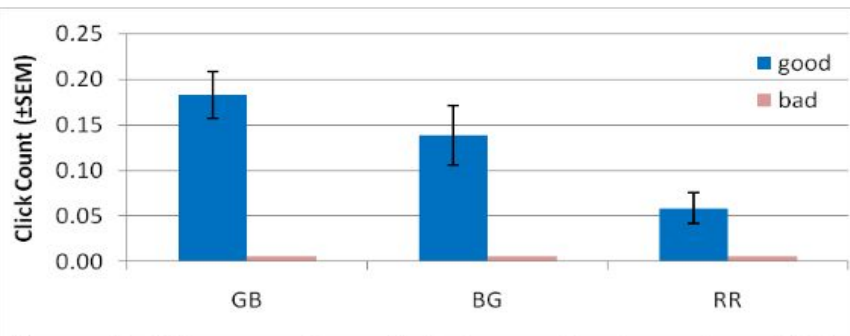
# The Good, The Bad, And The Fixation



**Figure 7: Comparison of mean fixation impact on SERP elements for navigational and informational tasks (left) and for SERPs displaying good or bad ads (right).**

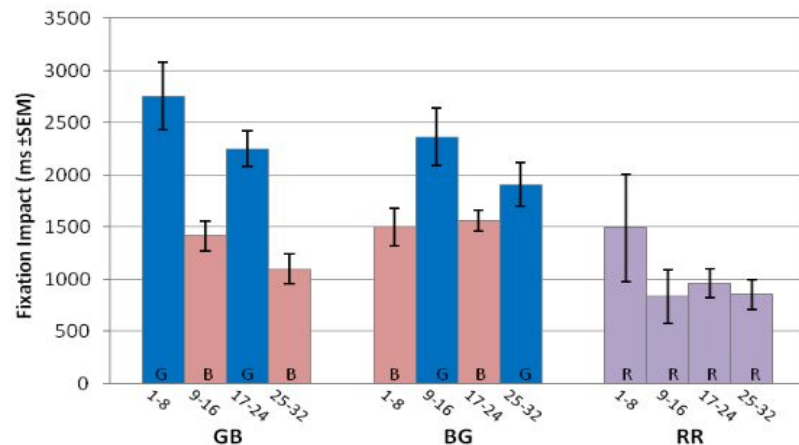


**Figure 8:** Mean fixation impact on the top ads  $fi(top\_ads)$  split by sequence of blocks (GB, BG, RR) and the quality of the displayed ads on the SERP (good / bad).



**Figure 9:** Mean number of clicks on the top ads split by sequence of blocks for good ads (there were no clicks on bad ads).

# Take The Ads Leave The Results



**Figure 10:** Mean fixation impact on the top ads  $fi(top\_ads)$  split by sequence of blocks (GB, BG, RR) and the block type (good, bad or random) of each block of 8 trials (see Figure 4).

# **Conclusion: Clickthrough Generates Revenue For Both Sites and Ads**

## **The Influence of Caption Features on Clickthrough Patterns in Web Search**

Clarke, Agichtein, Dumais & White Looked At Caption Features and Clickthrough Inversions

## **The Good, the Bad, and the Random: An Eye-Tracking Study of Ad Quality in Web Search**

Buscher, Dumais, & Cutrell Looked at Eye Movement And Add Quality for Clickthrough