

Beliefs & Biases in Web Search

Ryen White

Microsoft Research

ryenw@microsoft.com

Bias in IR and elsewhere

In IR, e.g.,

- Domain bias – People prefer particular Web domains
- Rank bias – People favor high-ranked results
- Caption bias – People prefer captions with certain terms

In psychology, e.g.,

- Anchoring-and-adjustment, confirmation, availability, etc.
- **All impact user behavior**
 - Opportunity to intersect psychology and IR

Our Interest in Biases

Bias can be observed in IR in situations where searchers seek or are presented with information that significantly deviates from the truth

More on the “truth” later...

Our Interest in Biases

User behavior

Search engine
behavior

Bias can be observed in situations where searchers seek or are presented with information that significantly deviates from the truth

More on the “truth” later...

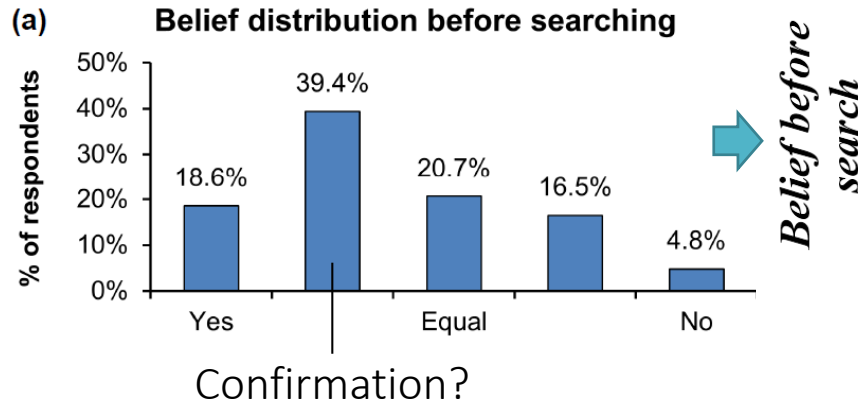
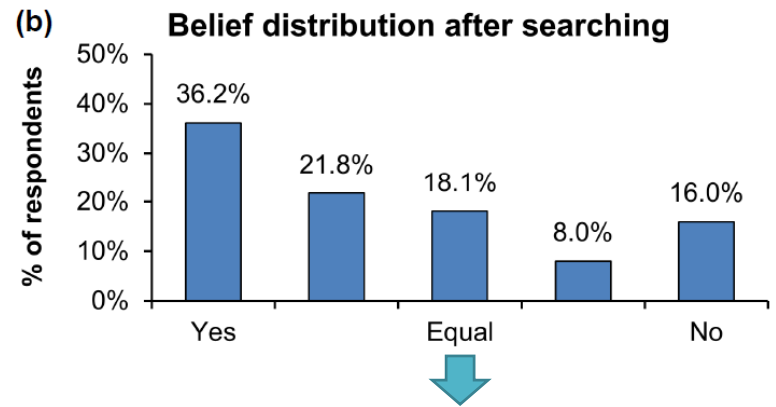
Outline for Remainder of Talk

- Initial Exploratory Questionnaire
- Log Analysis
- Labeling Content and Truth
- Findings
- Conclusions

Initial Exploratory Questionnaire

- Gain early insight into possible biases in search
- **Focus on Yes-No questions (answered with “Yes” or “No”)**
 - Simplicity: Answers along single dimension (Yes → No)
- Microsoft employees; recall recent Yes-No query (in last 2 weeks)
- Asked about belief beforehand and afterwards
 - Multi-point scale: **Yes** / Lean Yes / **Equal** / Lean No / **No**
- 200 respondents. Recalled questions such as:
 - *“Does chocolate contain caffeine?”*
 - *“Are shingles contagious?”*

Survey Results



	<i>Belief after search</i>				
	Yes		Equal		No
Yes	77.1%	8.6%	8.6%	2.9%	2.9%
Equal	23.1%	12.8%	43.6%	15.4%	5.1%
No	11.1%	—	11.1%	—	77.8%

Post-search belief given Pre-search belief

Two main findings:

1. Respondents kept strongly-held beliefs (Yes-Yes and No-No)
2. If Before = Equal, then 2x as likely to believe Yes after search

Motivated us to:

Further explore possible impact of biases on behavior and outcomes

Log-Based Study of Yes-No Queries

- Queries, clicks, and results from Bing logs (2 weeks)
- Mined yes-no questions: start with “can”, “is”, “does”, etc.
- Focused on health since it’s important and we could get truth
- Randomly selected set of 1000 yes-no health questions
 - Each issued by at least 10 users, same top 10, same captions
- Examples include:
 - *“Is congestive heart failure a heart attack?”* (answer = No)
 - *“Do food allergies make you tired?”* (answer = Yes)

Other Data Collected

- Yes-No Answer labels for captions and content of results
- Physician answers for the Yes-No questions

Answer Labeling

Example Caption Labels

Suggests **AFFIRMATIVE** answer (Yes only):

Question: [can i take l carnitine while pregnant]

Yes only

[Is it safe to take L-Carnitine while pregnant - The Q&A wiki](http://wiki.answers.com/Q/Is_it_safe_to_take_L-Carnitine_while_pregnant)

http://wiki.answers.com/Q/Is_it_safe_to_take_L-Carnitine_while_pregnant

Is l-carnitine safe to take while pregnant? yes. Is it safe to **take** zithromax **while pregnant?** yes it is safe to **take while pregnant.** A doctor would not prescribe it ...

Captions and result content

Crowdsourced (Clickworker.com)

3-5 judges/caption (consensus)

Task was to assign label of:

- **Yes** only
- **No** only
- **Both** (Yes and No)
- **Neither** (not Yes and not No)

Agreement on 96% of captions

- Performed similar labeling for each top 10 search results
 - Crowdsourced judges, agreement on 92% of pages

Answer Labeling

Example Caption Labels

Captions and result content

Crowdsourced (Clickworker.com)

3-5 judges/caption (consensus)

Task was to assign label of:

- **Yes** only
- **No** only
- **Both** (Yes and No)
- **Neither** (not Yes and not No)

Agreement on 96% of captions

- Performed similar labeling for each top 10 search results
 - Crowdsourced judges, agreement on 92% of pages

Suggests AFFIRMATIVE answer (Yes only):

Question: [can i take l carnitine while pregnant]

[Is it safe to take L-Carnitine while pregnant - The Q&A wiki](http://wiki.answers.com/Q/Is_it_safe_to_take_L-Carnitine_while_pregnant)

http://wiki.answers.com/Q/Is_it_safe_to_take_L-Carnitine_while_pregnant

Is l-carnitine safe to take while pregnant? yes. Is it safe to take zithromax while pregnant? yes it is safe to take while pregnant. A doctor would not prescribe it ...

Yes only

Suggests NEGATIVE answer (No only):

Question: [does robaxin show up on drug tests]

[Does robaxin show up on drug tests? | Answerbag](http://www.answerbag.com/q_view/1239474)

http://www.answerbag.com/q_view/1239474

Does robaxin show up on drug tests? no... More Questions. Additional questions in this category. Can you have a DUI & work at a school in Pennsylvania?

No only

Answer Labeling

Captions and result content

Crowdsourced (Clickworker.com)

3-5 judges/caption (consensus)

Task was to assign label of:

- **Yes** only
- **No** only
- **Both** (Yes and No)
- **Neither** (not Yes and not No)

Agreement on 96% of captions

- Performed similar labeling for each top 10 search results
 - Crowdsourced judges, agreement on 92% of pages

Example Caption Labels

Suggests AFFIRMATIVE answer (Yes only):

Question: [can i take l carnitine while pregnant]

[Is it safe to take L-Carnitine while pregnant - The Q&A wiki](http://wiki.answers.com/Q/Is_it_safe_to_take_L-Carnitine_while_pregnant)

http://wiki.answers.com/Q/Is_it_safe_to_take_L-Carnitine_while_pregnant

Is l-carnitine safe to take while pregnant? yes. Is it safe to **take** zithromax **while** pregnant? yes it is safe to **take while pregnant**. A doctor would not prescribe it ...

Yes only

Suggests NEGATIVE answer (No only):

Question: [does robaxin show up on drug tests]

[Does robaxin show up on drug tests? | Answerbag](http://www.answerbag.com/q_view/1239474)

http://www.answerbag.com/q_view/1239474

Does robaxin show up on drug tests? no... More Questions. Additional questions in this category. Can you have a DUI & work at a school in Pennsylvania?

No only

Suggests BOTH affirmative and negative:

Question: [is tooth a bone]

[Is tooth consider as a bone - The Q&A wiki](http://wiki.answers.com/Q/Is_tooth_consider_as_a_bone)

http://wiki.answers.com/Q/Is_tooth_consider_as_a_bone

What does the **bone** in the **tooth** do? It helps u chew. **Is a tooth a bone? Yes. Is your tooth a bone? No, teeth are not bones.** Is the "skin" lining your stomach skin?

Both

Answer Labeling

Example Caption Labels

Captions and result content

Crowdsourced (Clickworker.com)

3-5 judges/caption (consensus)

Task was to assign label of:

- Yes only
- No only
- Both (Yes and No)
- Neither (not Yes and not No)

Agreement on 96% of captions

- Performed similar labeling for each top 10 search results
 - Crowdsourced judges, agreement on 92% of pages

Suggests AFFIRMATIVE answer (Yes only):

Question: [can i take l carnitine while pregnant]

[Is it safe to take L-Carnitine while pregnant - The Q&A wiki](http://wiki.answers.com/Q/Is_it_safe_to_take_L-Carnitine_while_pregnant)

http://wiki.answers.com/Q/Is_it_safe_to_take_L-Carnitine_while_pregnant

Is l-carnitine safe to take while pregnant? yes. Is it safe to **take** zithromax **while pregnant?** yes it is safe to **take while pregnant.** A doctor would not prescribe it ...

Yes only

Suggests NEGATIVE answer (No only):

Question: [does robaxin show up on drug tests]

[Does robaxin show up on drug tests? | Answerbag](http://www.answerbag.com/q_view/1239474)

http://www.answerbag.com/q_view/1239474

Does robaxin show up on drug tests? no... More Questions. Additional questions in this category. Can you have a DUI & work at a school in Pennsylvania?

No only

Suggests BOTH affirmative and negative:

Question: [is tooth a bone]

[Is tooth consider as a bone - The Q&A wiki](http://wiki.answers.com/Q/Is_tooth_consider_as_a_bone)

http://wiki.answers.com/Q/Is_tooth_consider_as_a_bone

What does the **bone** in the **tooth** do? It helps u chew. **Is a tooth a bone? Yes. Is your tooth a bone? No, teeth are not bones.** Is the "skin" lining your stomach skin?

Both

Suggests NEITHER affirmative nor negative:

Question: [does crestor cause bloating]

[Does Crestor Cause Bloating? - HealthCentral](http://www.healthcentral.com/cholesterol/h/does-crestor-cause-bloating.html)

<http://www.healthcentral.com/cholesterol/h/does-crestor-cause-bloating.html>

Everything you need to know about **does crestor cause bloating**, including common uses, side effects, interactions and risks.

Neither

Physician Answers

- Two physicians reviewed the 1000 questions and gave answers
 - Inc. **50/50** = need more info, **Don't know** = really unsure
- Agreement between physicians on Yes-No was 84% ($\kappa=0.668$)

		<i>Physician 2</i>				<i>Total</i>
		<i>Yes</i>	<i>No</i>	<i>50/50</i>	<i>Don't know</i>	
<i>Physician 1</i>	<i>Yes</i>	38.8%	8.2%	3.7%	0.5%	51.2%
	<i>No</i>	5.7%	31.5%	1.2%	0.2%	38.5%
	<i>50/50</i>	1.8%	2.0%	1.3%	0.0%	5.0%
	<i>Don't know</i>	1.3%	3.1%	0.2%	0.7%	5.3%
	<i>Total</i>	47.5%	44.8%	6.3%	1.5%	100.0%

- Focused on the **680 questions** where both agreed Yes or No
- Distribution: **55% Yes** and **45% No** (used as **TRUTH** in our study)

Using Physician Answers as Truth

Used consensus physician answers as truth in three ways:

- How closely does distribution of results match the truth?
- How closely does interaction behavior match the truth?
- How closely do answers that people reach match the truth?

Bias = Distributions significantly differ from 55-45 Yes-No base rates

Taking Stock of Our Data

- We have:
 - 680 Yes-No health questions from search logs
 - Ground truth for each q via physicians' consensus judgments
- For each question we have:
 - *HTML content of top 10 search results, plus:*
 - Caption labels for Yes/No/Both/Neither
 - Result labels for Yes/No/Both/Neither
 - *Clickthrough behavior from logs*

Analysis

- Three directions for analysis:
 - Study **ranking of results** with Yes-No content
 - Study **user behavior** w.r.t. Yes-No content
 - Study **answer accuracy** for Yes-No questions

Result Ranking

- Volume of Yes-No content in the results

Percentage of captions or results with answer

<i>Source</i>	<i>Yes only</i>	<i>No only</i>	<i>Both</i>	<i>Neither</i>
Caption	28.7%	8.4%	2.7%	60.2%
Result	35.0%	12.7%	6.3%	41.0%

→ More Yes content in top-10 than No content

- Relative ranking of top Yes-No content when both in top 10

Percentage of SERPs where top *yes* caption or result appears above (nearer the top of the ranking than) the top *no*

<i>Source</i>	<i>Yes above No</i>	<i>No above Yes</i>
Caption	65.1%	34.9%
Result	62.4%	37.6%

→ Yes content ranked above No more often (when both shown)

User Behavior (Clickthrough rate)

- Studied **clickthrough rates** on captions containing answers
- Controlled for rank by just considering top result ($r=1$)

SERP click likelihoods for different captions given variations in answer presence in SERPs/captions, and rank

<i>Condition(s)</i>	<i>All</i>	
$SERP_Y$	80.0%	3-4x as likely to click on captions with Yes content, even though TRUTH = 55% Yes / 45% No
$SERP_N$	75.9%	
$SERP_{BOTH}, Caption_Y$	45.6%	
$SERP_{BOTH}, Caption_N$	14.2%	
$Caption_Y$	41.1%	
$Caption_N$	16.3%	
$Caption_{Y,r=1}$	47.4%	
$Caption_{N,r=1}$	12.6%	

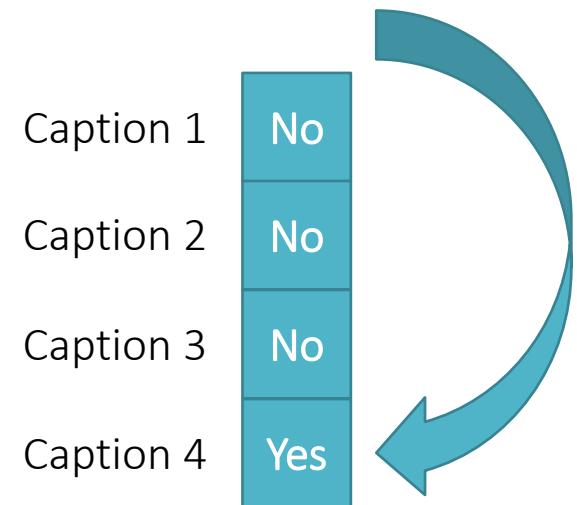
Just considering top search result

User Behavior (Result skipping)

- Studied result **skipping** behavior
- Frequency with which people skipped caption w/answer to click other caption

Distribution of clicks and skips by answer

<i>Click</i>	<i>Skip</i>	
	<i>Yes only</i>	<i>No only</i>
<i>Yes only</i>	33.3%	41.5%
<i>No only</i>	8.5%	16.7%



- Users more likely (4x) to skip No to click Yes than vice versa

Answer Accuracy

- Examined accuracy of the top search result, as well as first click and last click in session

<i>Answer defn.</i>	<i>All</i>	<i>Physician Answer</i>	
		<i>Yes</i>	<i>No</i>
Top result	45.0%	57.1%	22.9%
First click	50.0%	59.1%	27.9%
Last click	52.3%	66.2%	29.4%

- Findings show:
 1. Top result accurate only 45% of time, less when truth is No
 2. Users improve accuracy, but only slightly (limited by top 10)

Summary of Main Findings

We observed:

1. Engines more likely to rank Yes above No, and return more Yes
2. People much more likely to click on Yes than No, even when control for availability and rank position
3. Engine had wrong answer @ top rank for half of questions*
* Given that answer present at top position (~80% of queries)

- Caveats:
 - Findings for our particular set of Yes-No health questions
 - More work needed to validate with other question sets, domains beyond health, etc.

Discussion

- Possible causes for observed bias include:
 - Search engines use behavior (hurt by common misconceptions)
 - Ranking algorithms consider query match

e.g., for query: [*can acid reflux cause back pain?*]:

Yes docs w/ “Acid reflux can cause back pain” better match (6 of 6 terms) than **No** docs w/ “Acid reflux cannot cause back pain” (5 of 6 terms) **missing from query**

Conclusions

- Studied potential bias in user behavior and outcomes
- Showed effects on both from search engines
- 2% of queries are Yes-No questions; Searchers want answers!!
- To get users to *accurate* answers, engines should consider truth
- Future directions:
 - Study availability of Yes-No content online; Move beyond Yes-No
 - Consider how truth should be determined and used in ranking
 - Follow-up user studies