

BubbleView: an interface for crowdsourcing image importance maps and tracking visual attention



**Nam Wook
Kim***



**Zoya
Bylinskii***



Michelle
Borkin



Krzysztof
Gajos



Aude
Oliva



Fredo
Durand



Hanspeter
Pfister



HARVARD
John A. Paulson
School of Engineering
and Applied Sciences



Northeastern University
College of Computer and Information Science

Eye-tracking for capturing human visual attention



Eye-tracking for capturing human visual attention

in-lab experiment

tedious calibration

specialized hardware



Eye-tracking for capturing human visual attention

in-lab experiment

Difficult to scale up data collection to more than a few participants

tedious calibration

specialized hardware

0 clicks

0 characters

Describe the image in as much detail as possible...

Bubble View

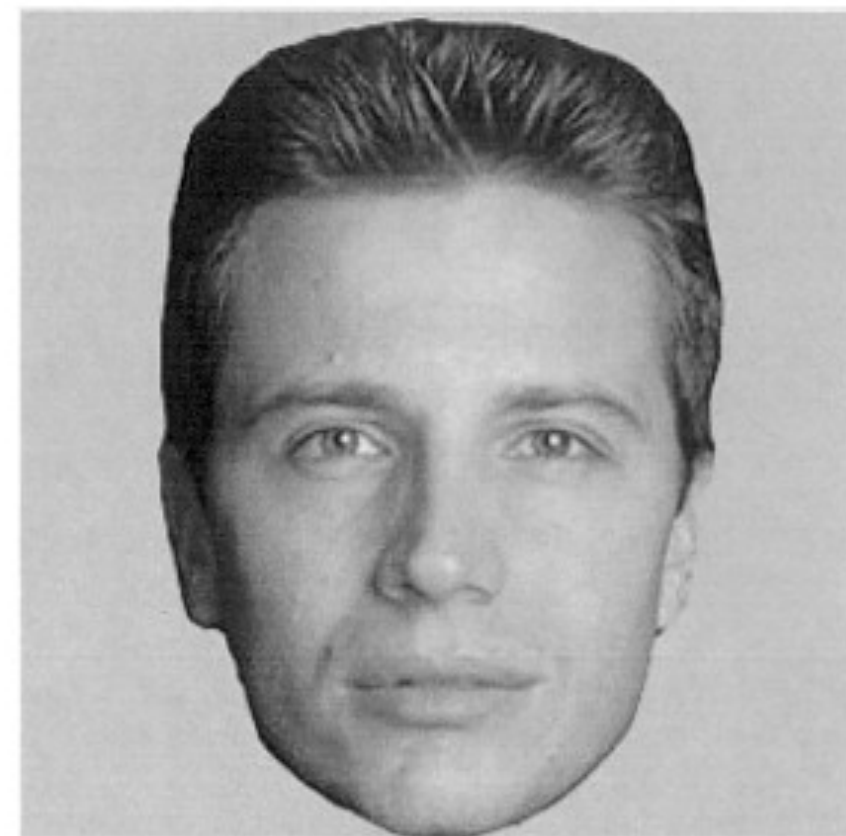
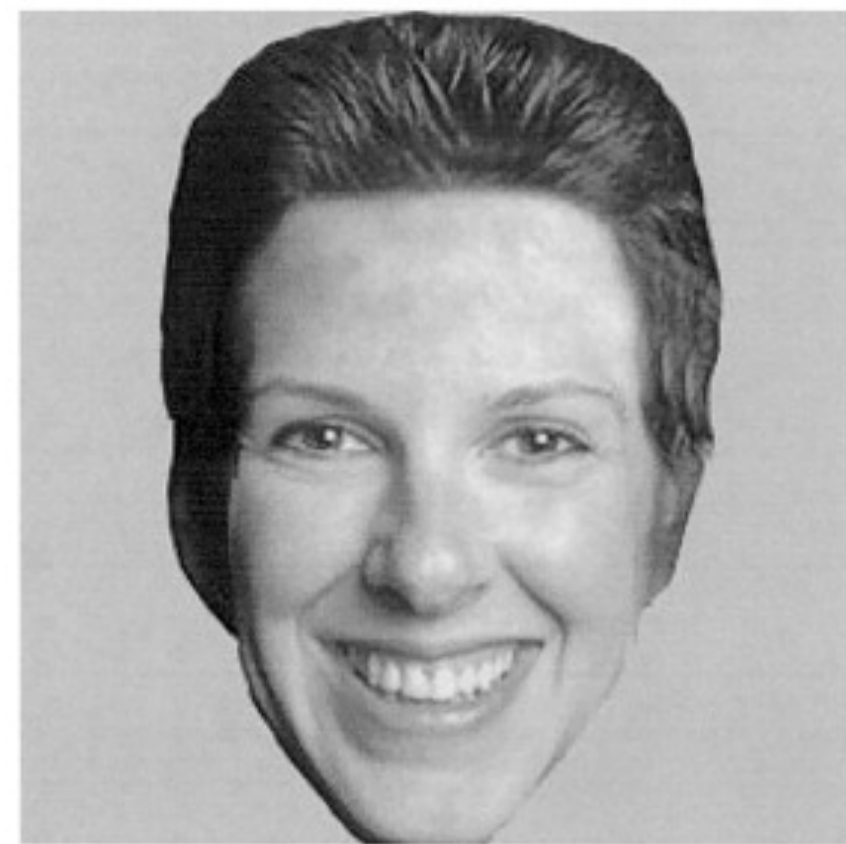
An alternative for eye tracking using **discrete mouse clicks** to measure which information people **consciously** choose to examine.

» 2x

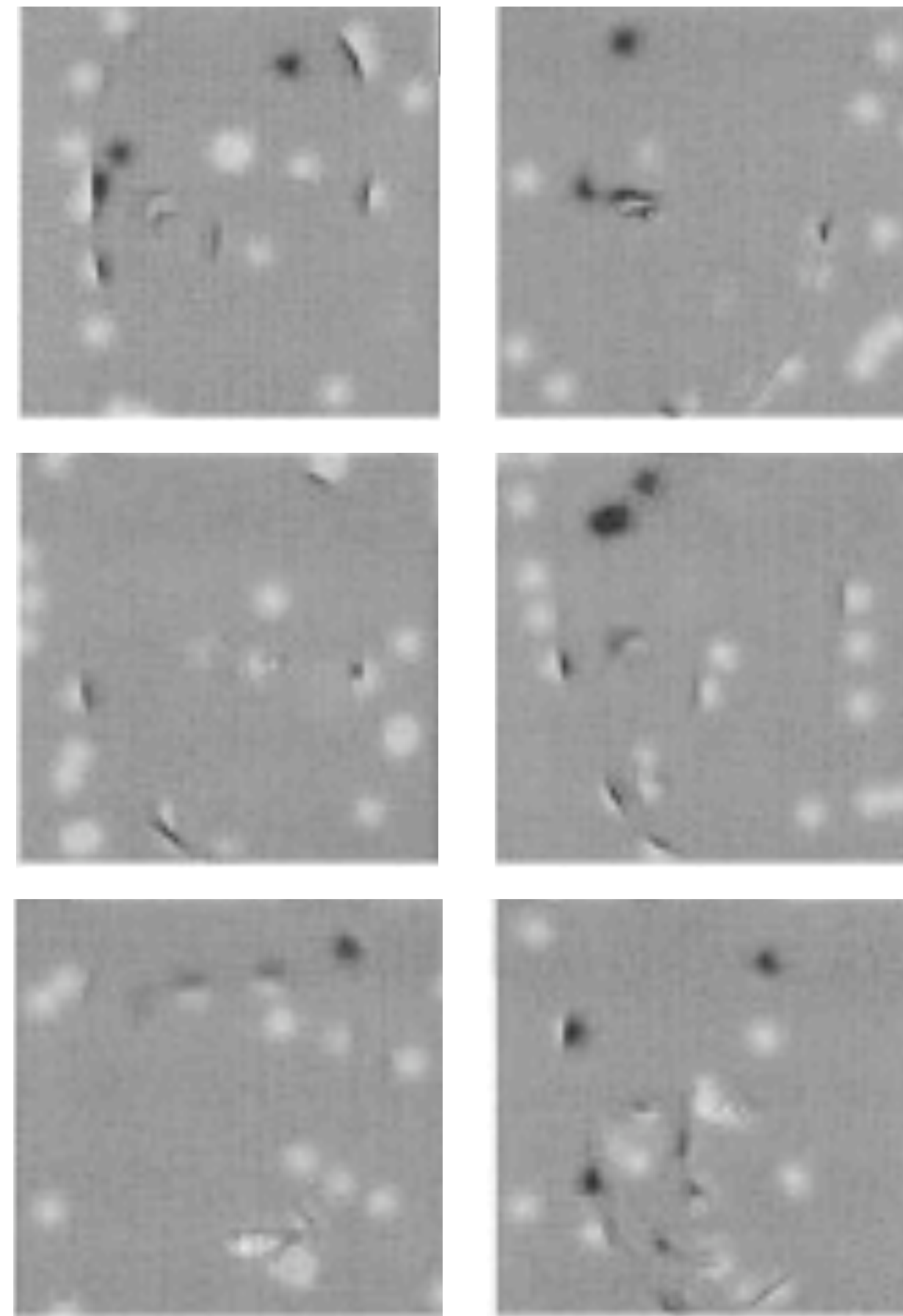
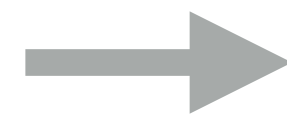
NEXT

Inspiration: Bubbles

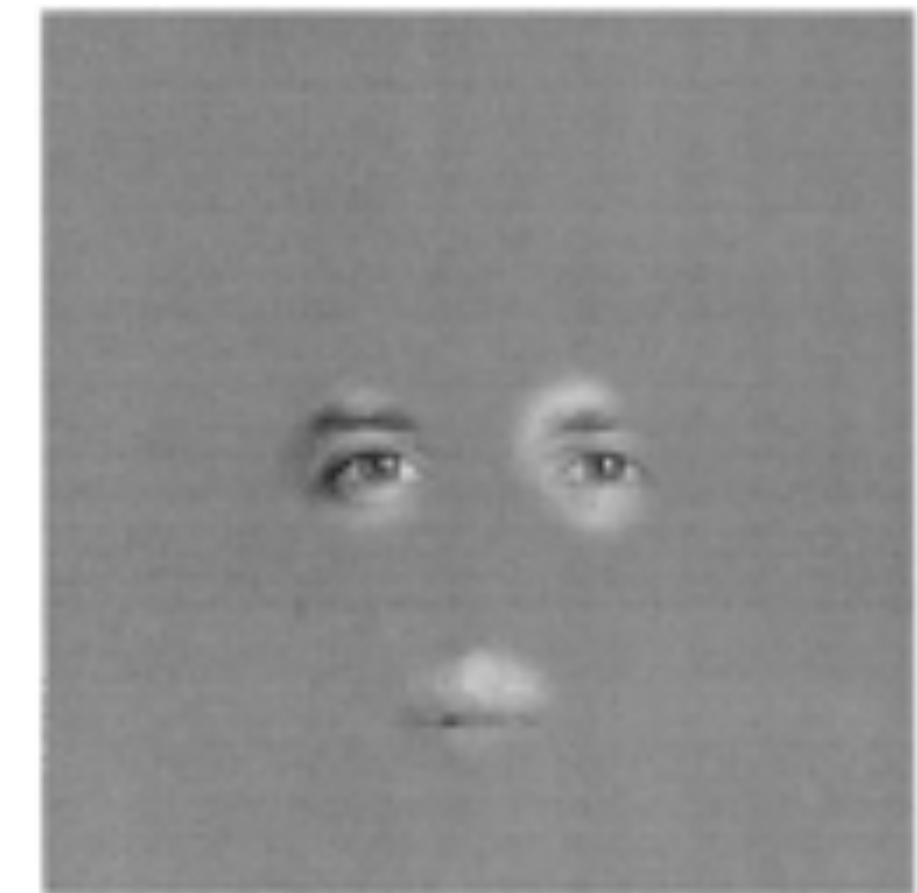
[Gosselin & Schyns, 2001]



Face stimuli



Punctured by bubbles

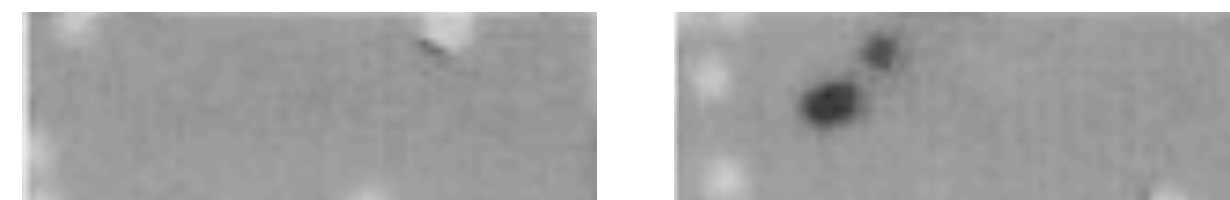
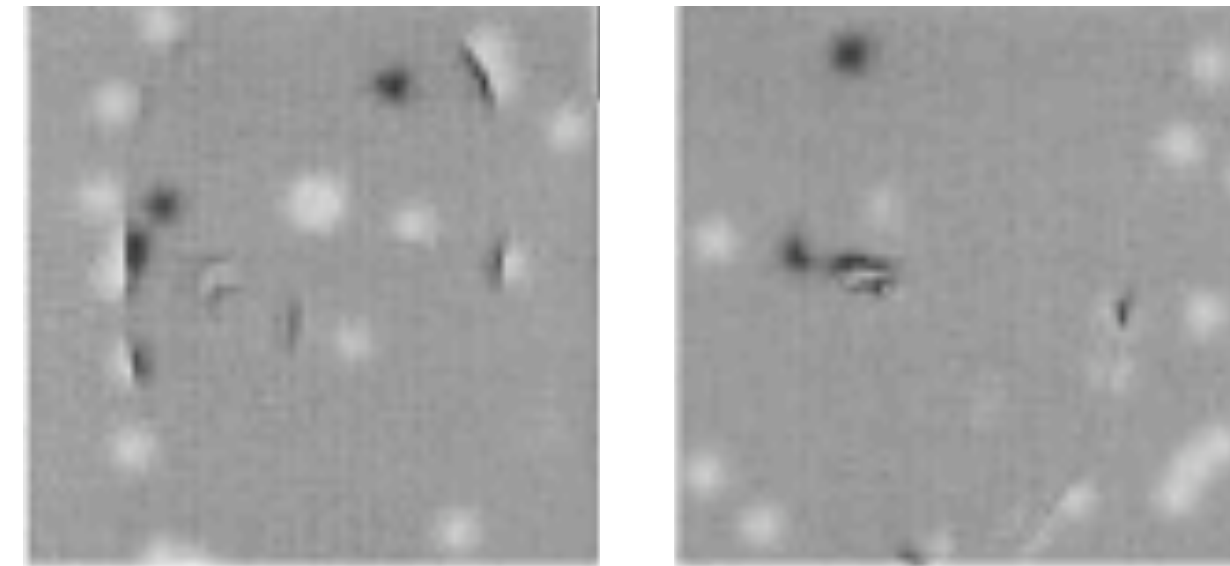
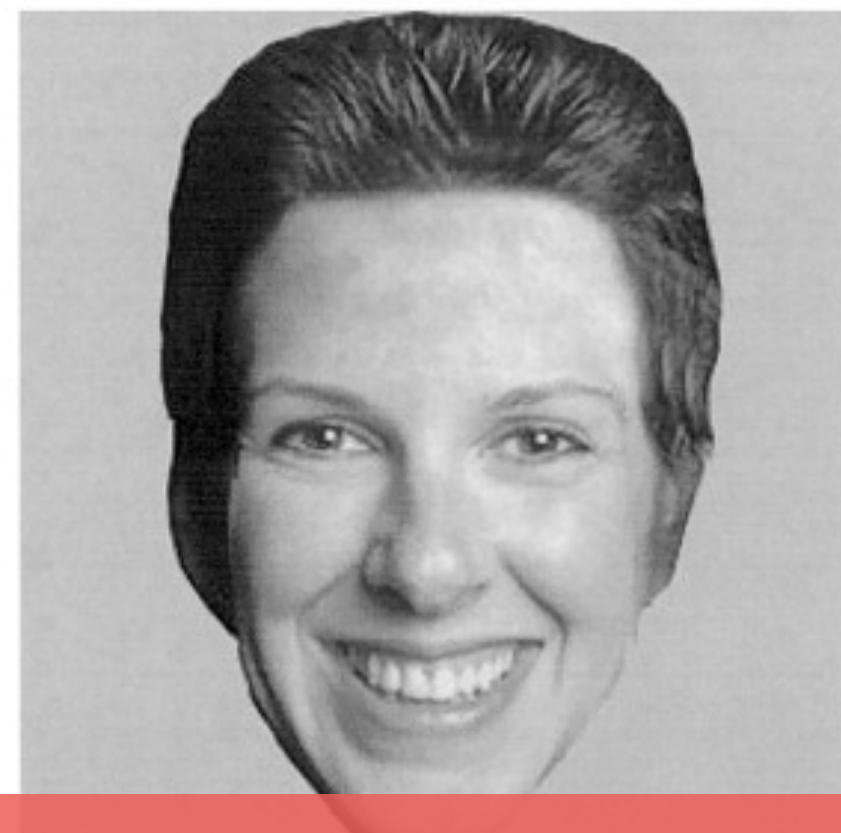


The **eyes** and **mouth**

Gender categorization

Inspiration: Bubbles

[Gosselin & Schyns, 2001]



BubbleView generalizes this idea to allow users to control where they want to look.

The eyes and mouth

Face stimuli

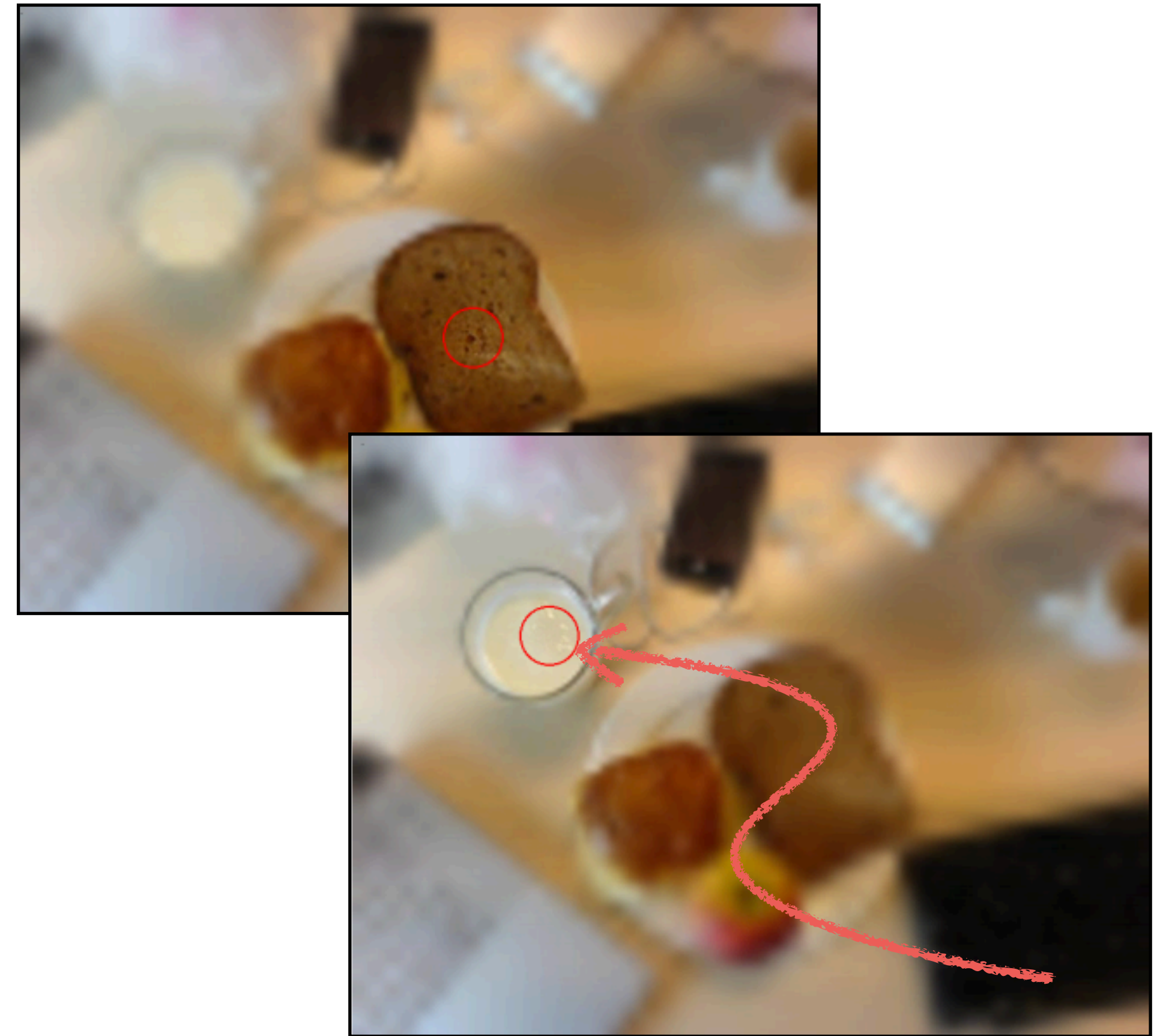
Punctured by bubbles

Gender categorization

Cursor-Based Attention Tracking

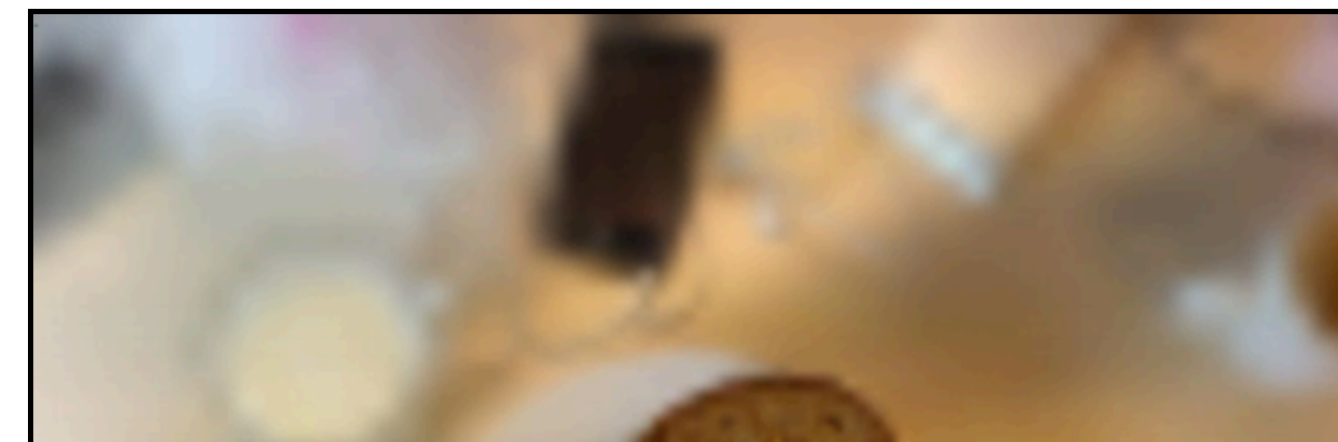
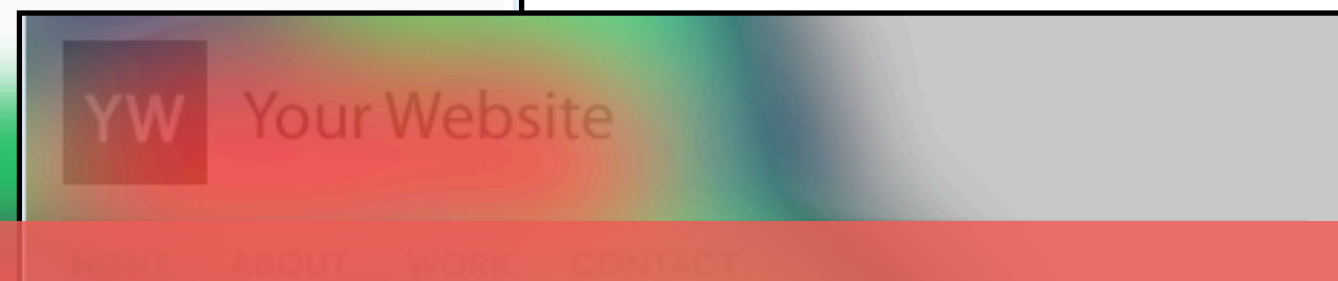
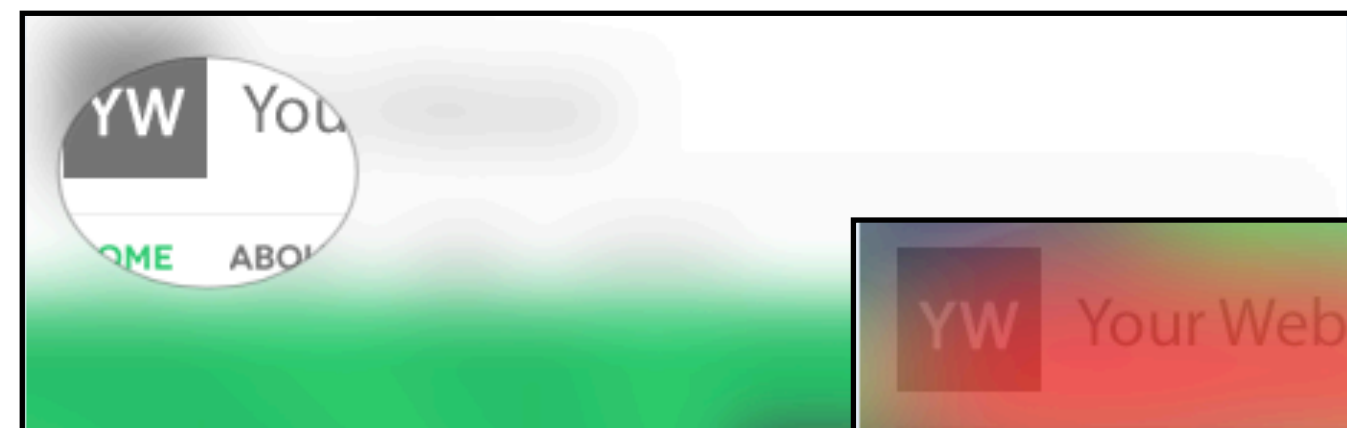


[Schulte- Mecklenbeck et al. 2011]

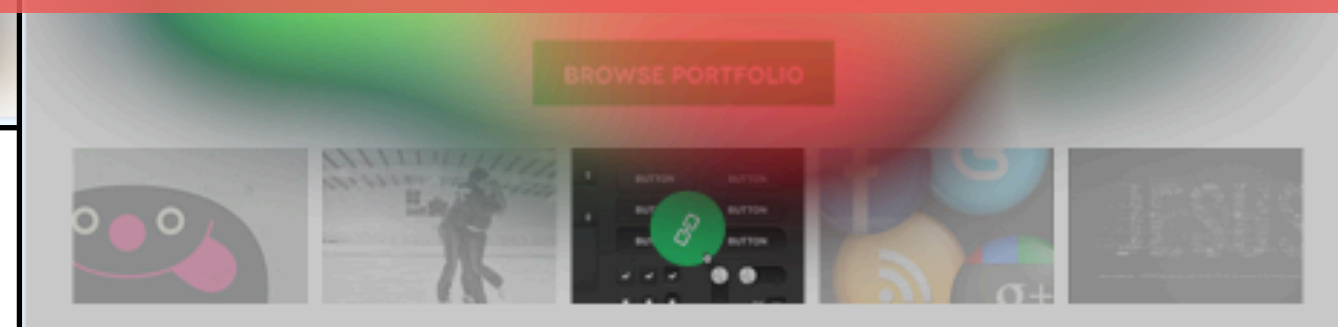


[Jiang et al. 2015]

Cursor-Based Attention Tracking



Discrete **clicks** instead of continuous **movements**
to explicitly record points of interest



[Schulte- Mecklenbeck et al. 2011]

[Jiang et al. 2015]

Cursor-Based Attention Tracking

We systematically evaluate cursor-based tracking under different parameters and task settings.

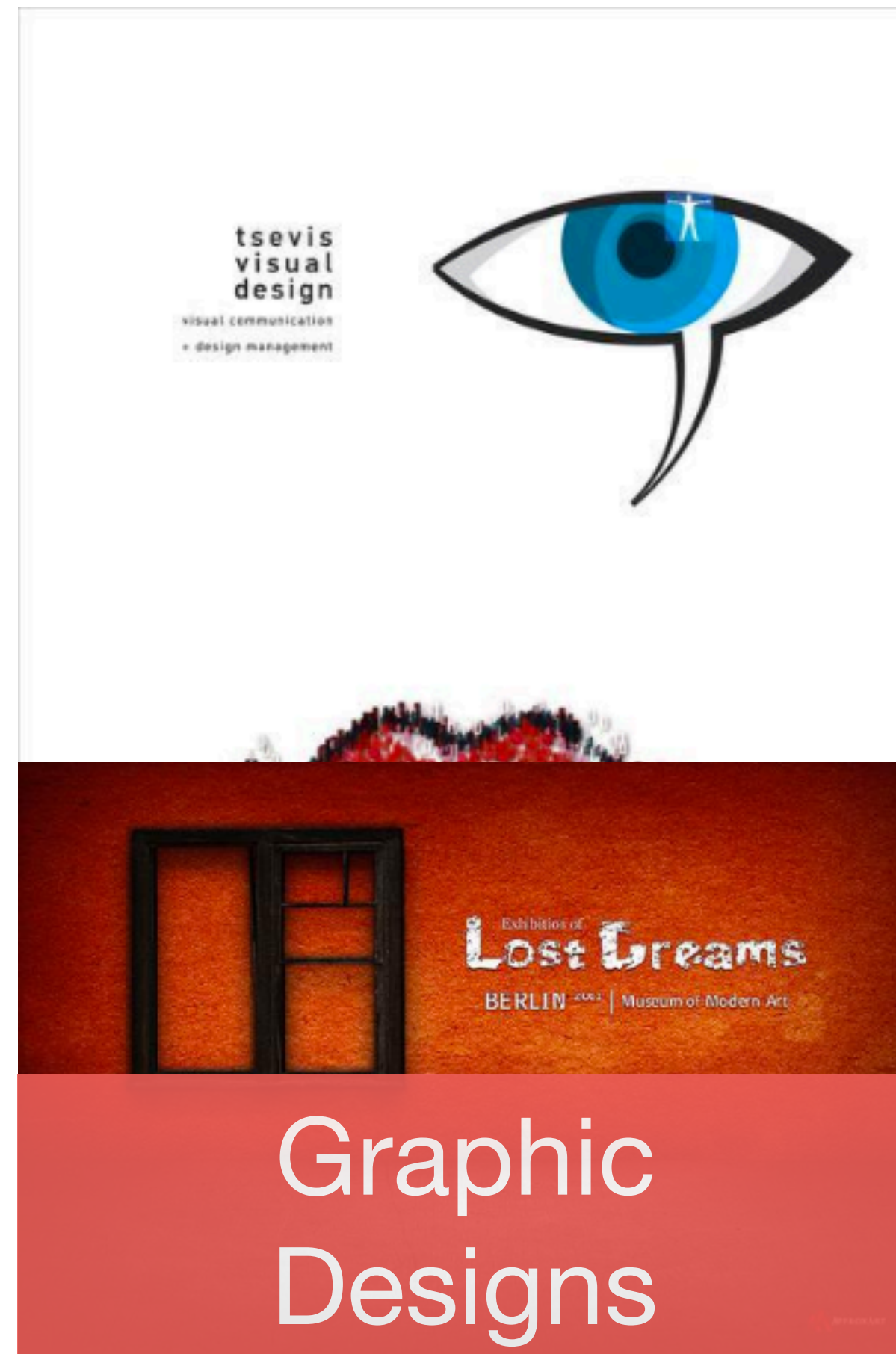
[Schulte- Mecklenbeck et al. 2011]

[Jiang et al. 2015]

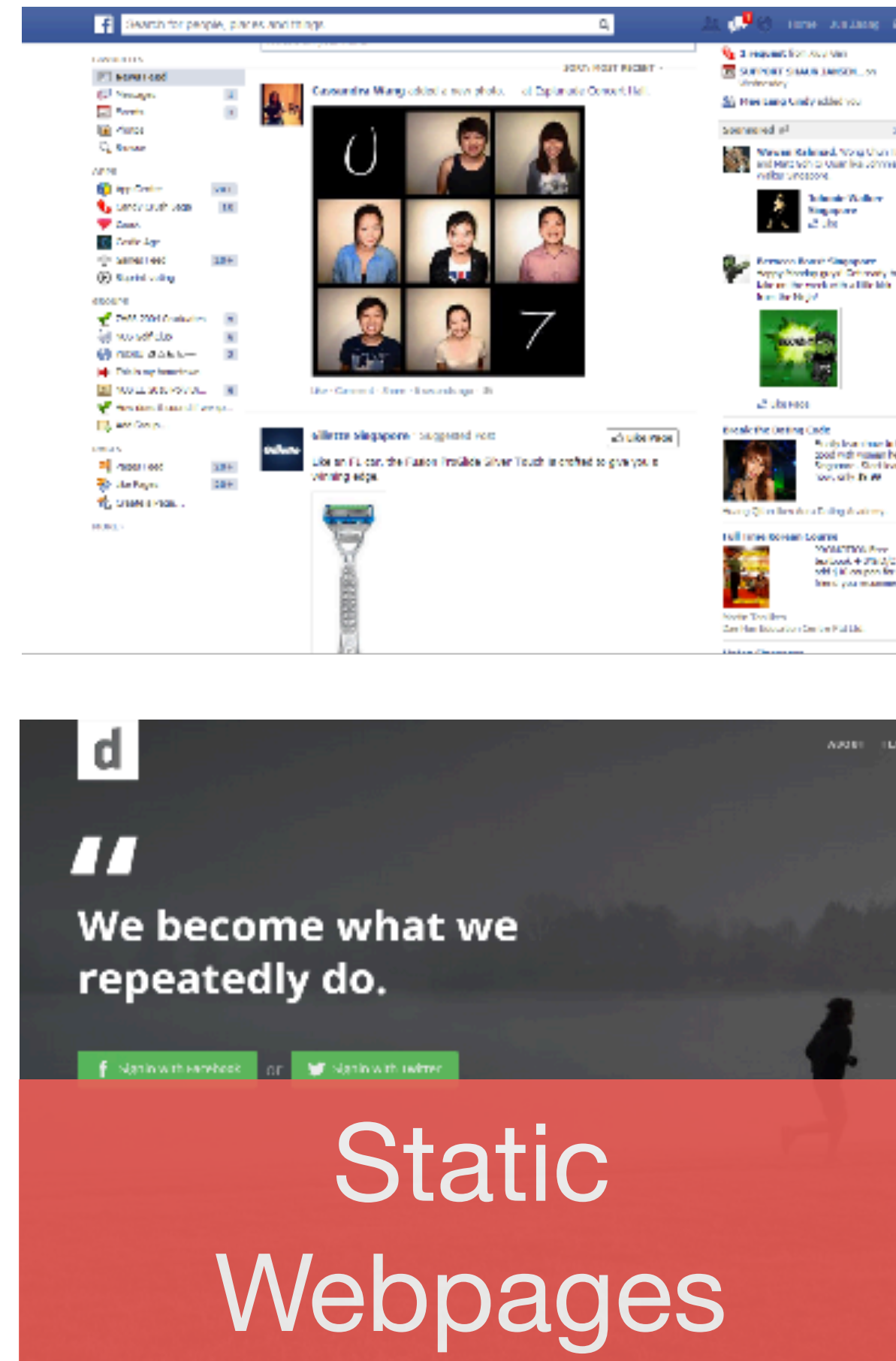
Evaluated on Various Image Types



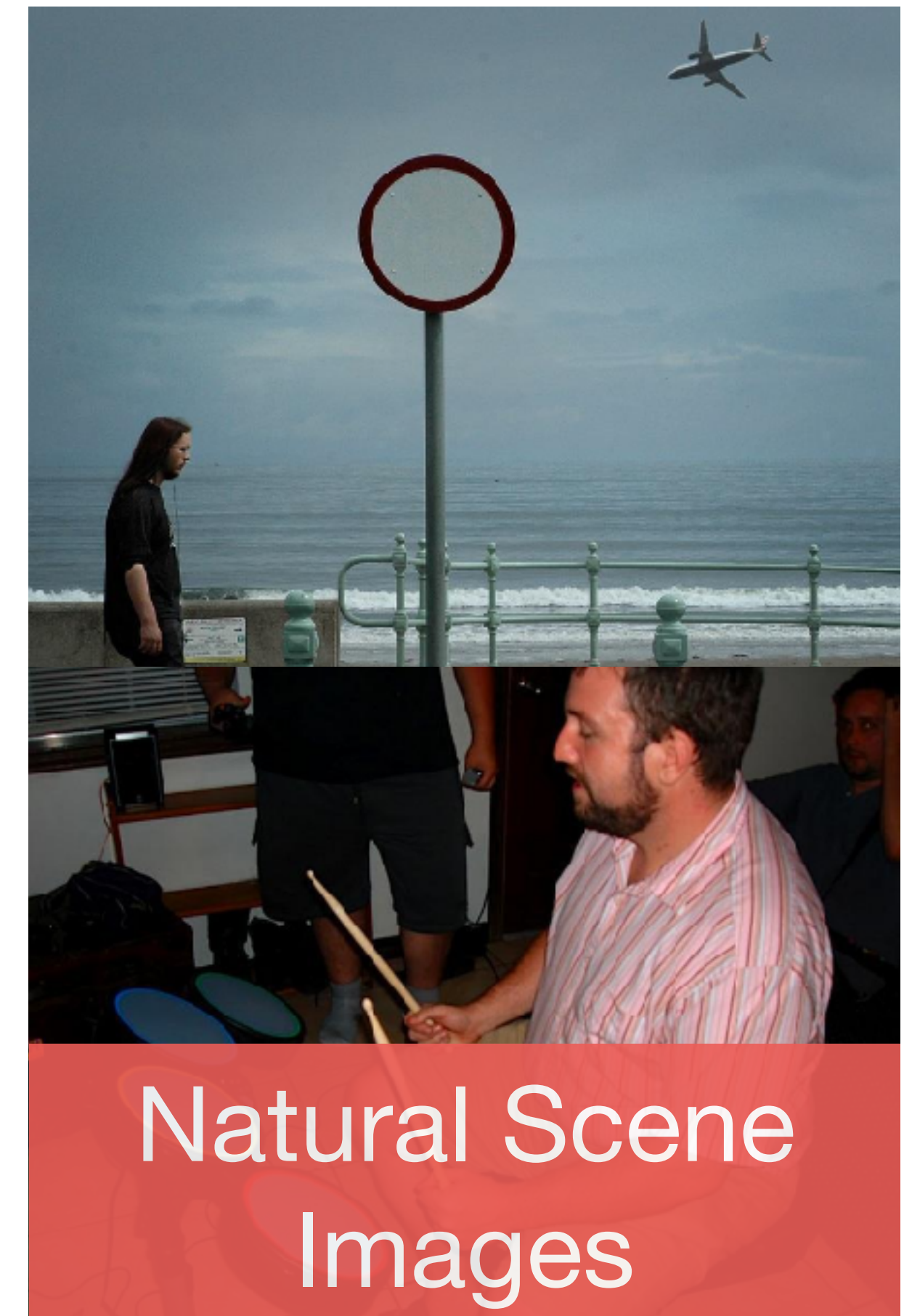
MASSVIS
[Borkin et al. 2016]



GDI
[O'Donovan et al. 2014]



FIWI
[Shen and Zhao 2014]



OSIE
[Xu et al. 2014]

Evaluation Configuration



Information Visualizations

Describe
(unlimited time)

Bubble radius
(16,24,32,40)

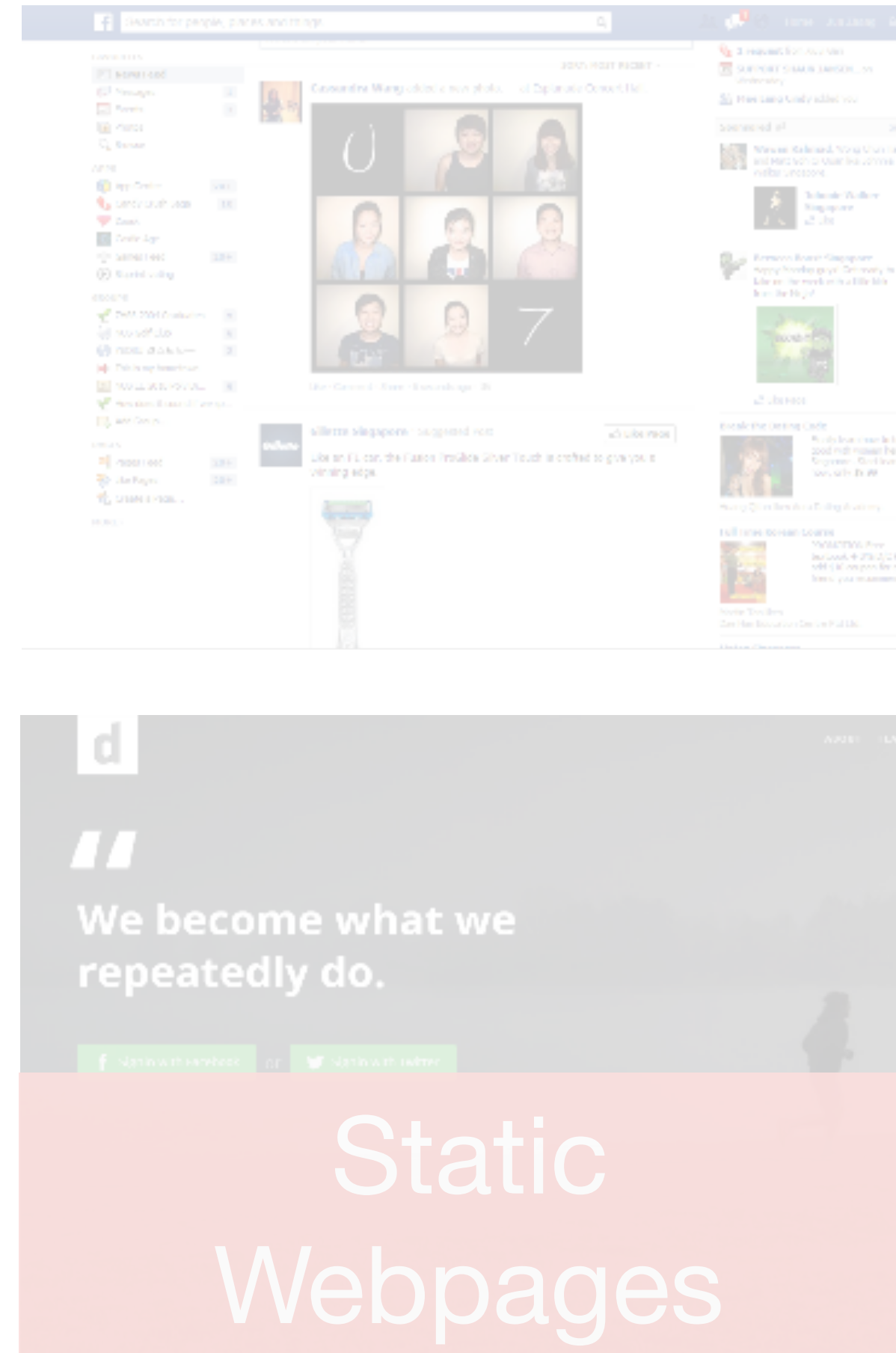
vs Fixations

The image shows a red-tinted background with various data visualizations. At the top, there's a bar chart titled 'Cities' showing percentage change. Below it, a text snippet 'Bond Boom' is visible. At the bottom, a bar chart shows data from 2000 to 2012.



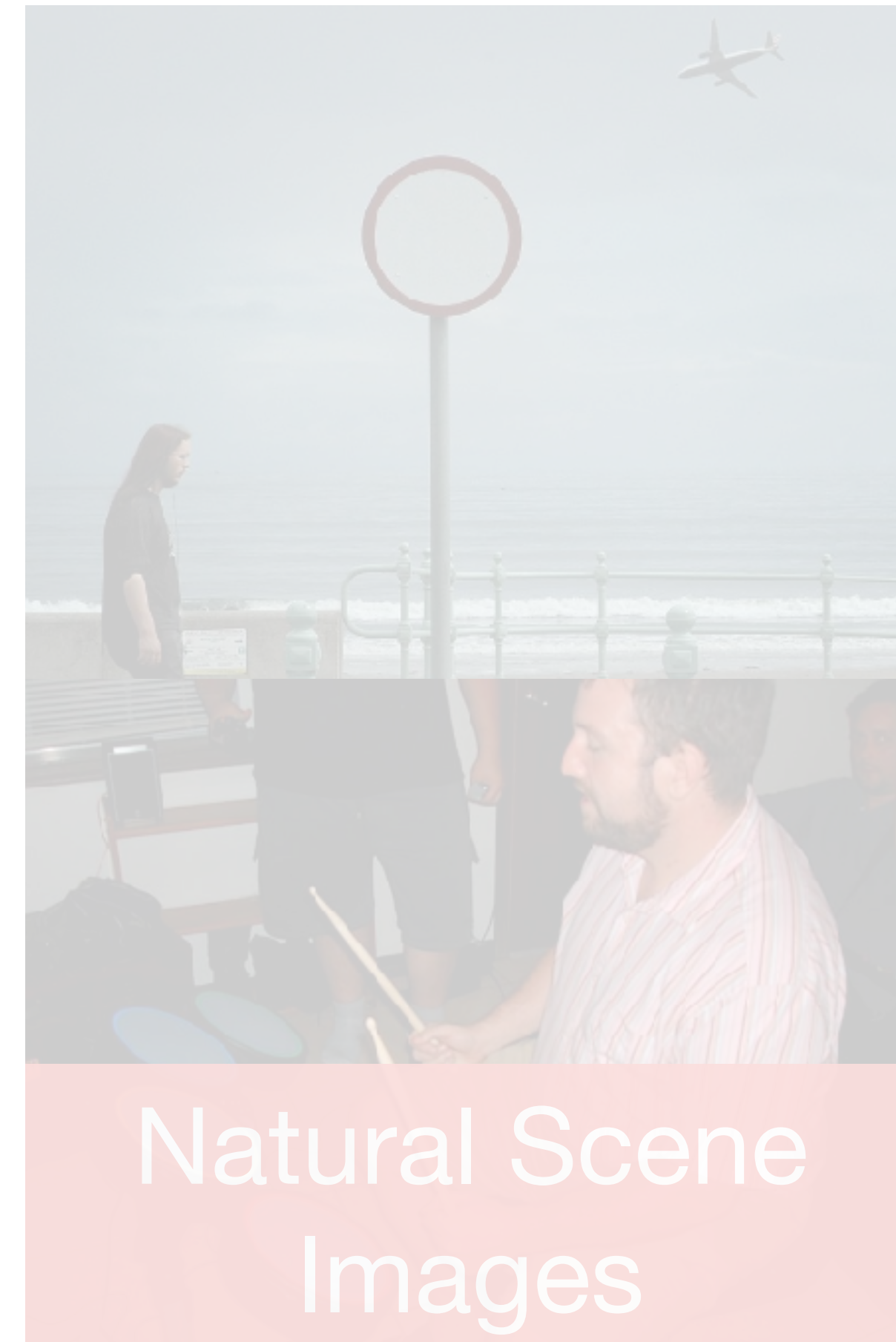
Graphic Designs

The image displays two graphic design examples. The top one is a logo for 'tsevis visual design' featuring a stylized blue eye with a white pupil and a white speech bubble tail. The bottom one is a poster for an exhibition titled 'Lost Dreams' at the Museum of Modern Art in Berlin, showing a building facade with a window.



Static Webpages

The image shows two static webpage examples. The top one is a Facebook profile page for 'Catherine Wang' with a grid of profile pictures. The bottom one is a webpage with a dark background and a quote: 'We become what we repeatedly do.' with social media sharing buttons.



Natural Scene Images

The image shows two natural scene photographs. The top one is a person standing on a pier looking out at the ocean with a circular signpost in the foreground. The bottom one is a man in a striped shirt playing a drum set.

Description Task

Click and describe the image



0 clicks 0 characters

Describe the image in as much detail as possible...

Unlimited time +
150 minimum characters

NEXT

Varied Bubble Sizes

How does bubble radius size affect performance?



16 pixels



24 pixels



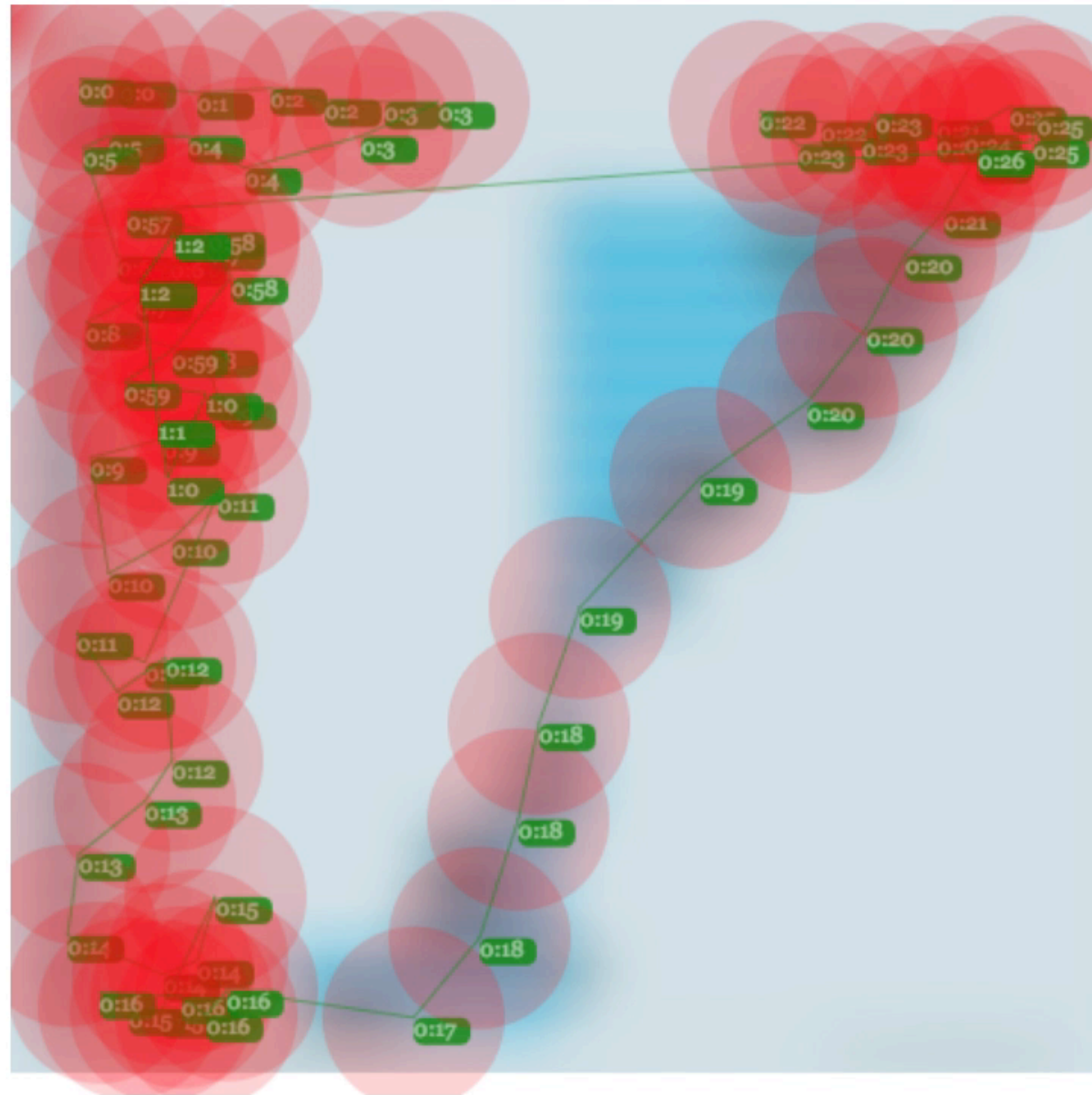
32 pixels



40 pixels

Collected Data

See Original See Bubbles



Time Slider (75/75)

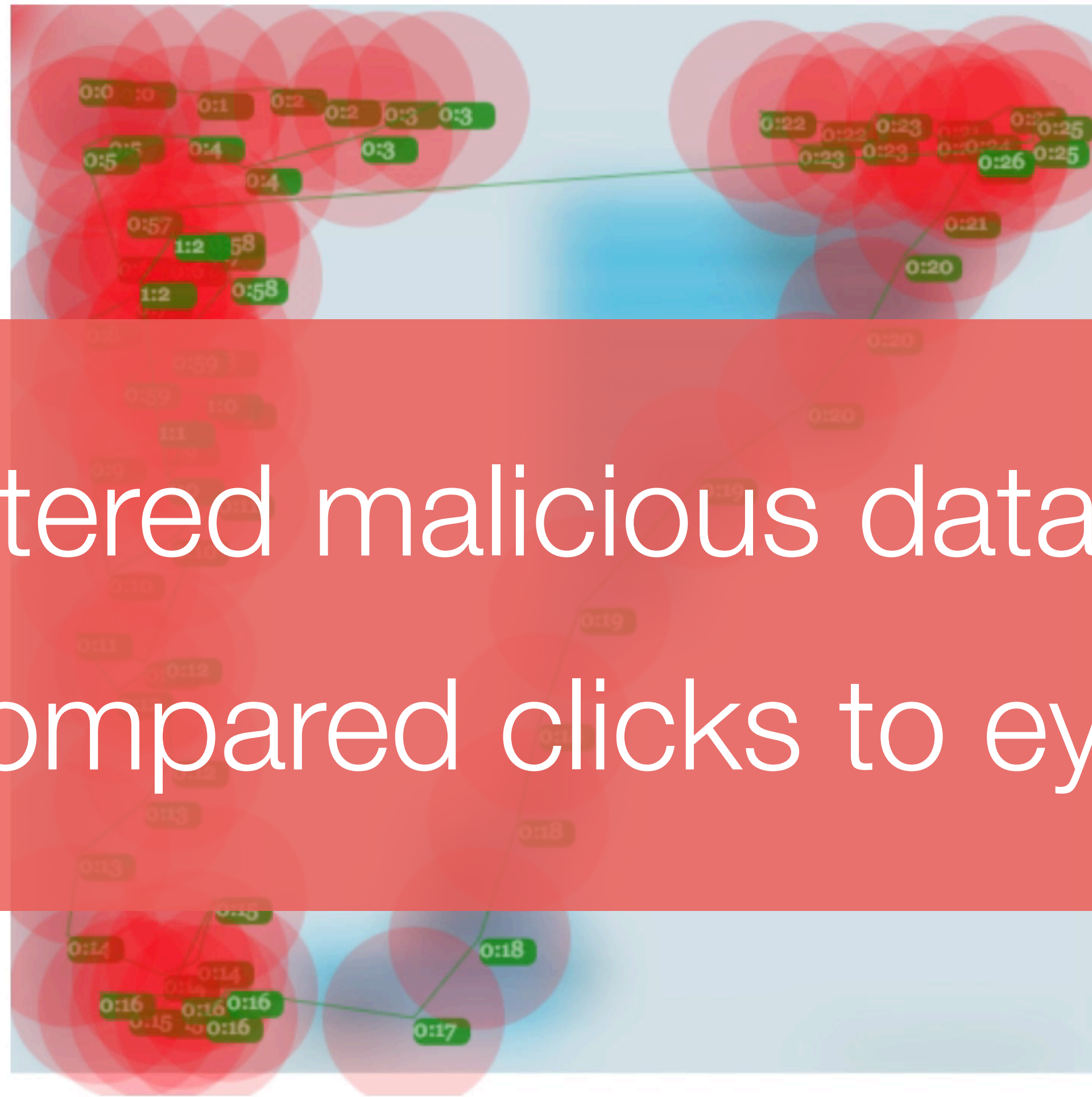
Description Change:

This chart includes the percentage change on the previous year forecast of income versus employment for the year 2010. Cities include Istanbul, Lima, Singapore, Shanghai, Mumbai, Rio de Janeiro.

Clicks &
Description changes over time.

Collected Data

See Original See Bubbles



Time Slider (75/75)

Description Change:

This chart includes the percentage change on the previous year forecast of income versus employment for the year 2010. Cities include Istanbul, Lima, Singapore, Shanghai, Mumbai, Rio de Janeiro.

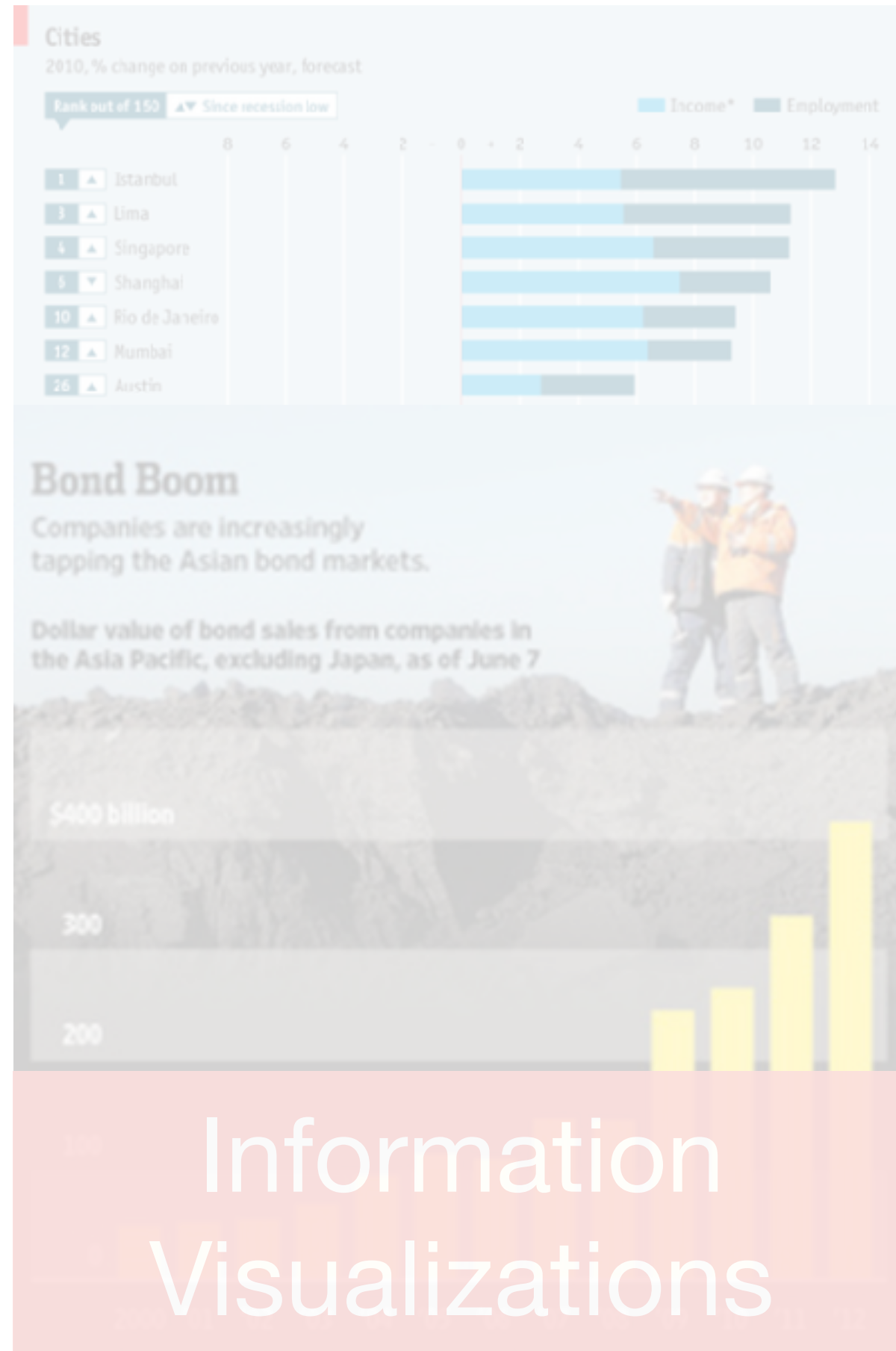
Filtered malicious data &
Compared clicks to eye-fixations

Clicks &
Description changes over time.



MASSVIS
[Borkin et al. 2016]

Evaluation Configuration



Graphic Designs

Free-view
(10 sec)

vs Annotations

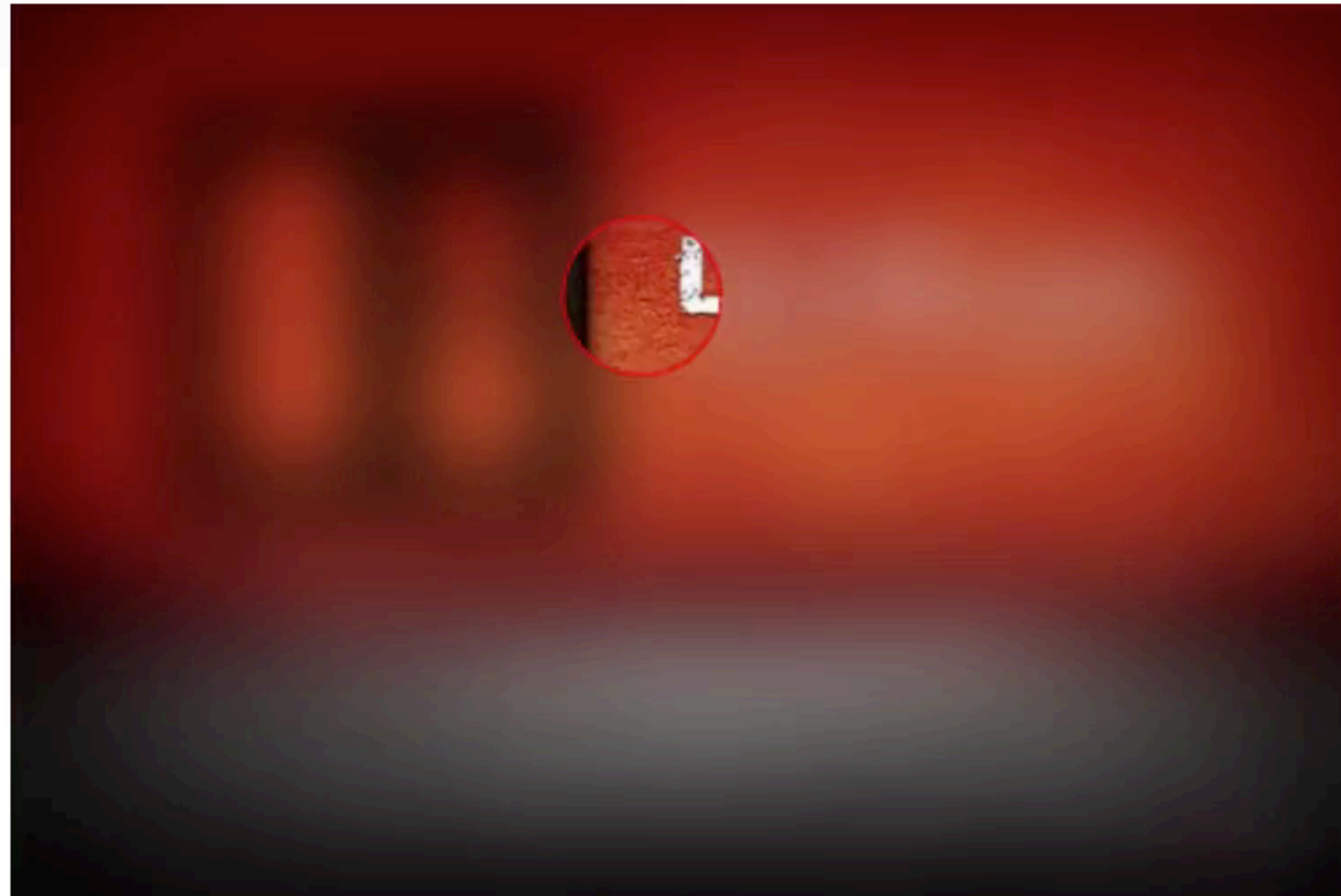
Static Webpages

Natural Scene Images

Much less informational content in graphic design images

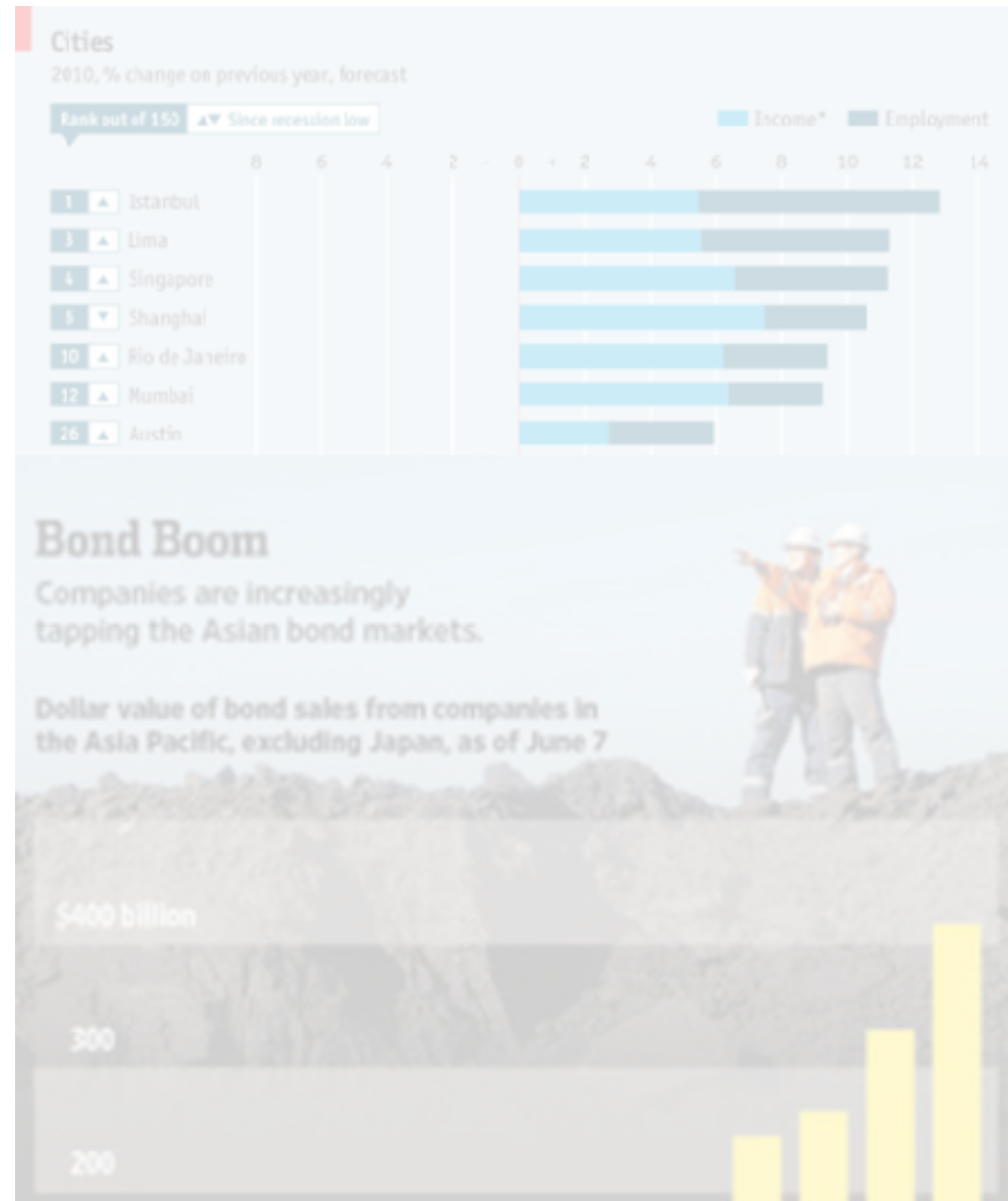
Free-Viewing Task

Click anywhere you want to look.



10 seconds of viewing
No description required

Evaluation Configuration



Information Visualizations



Graphic Designs

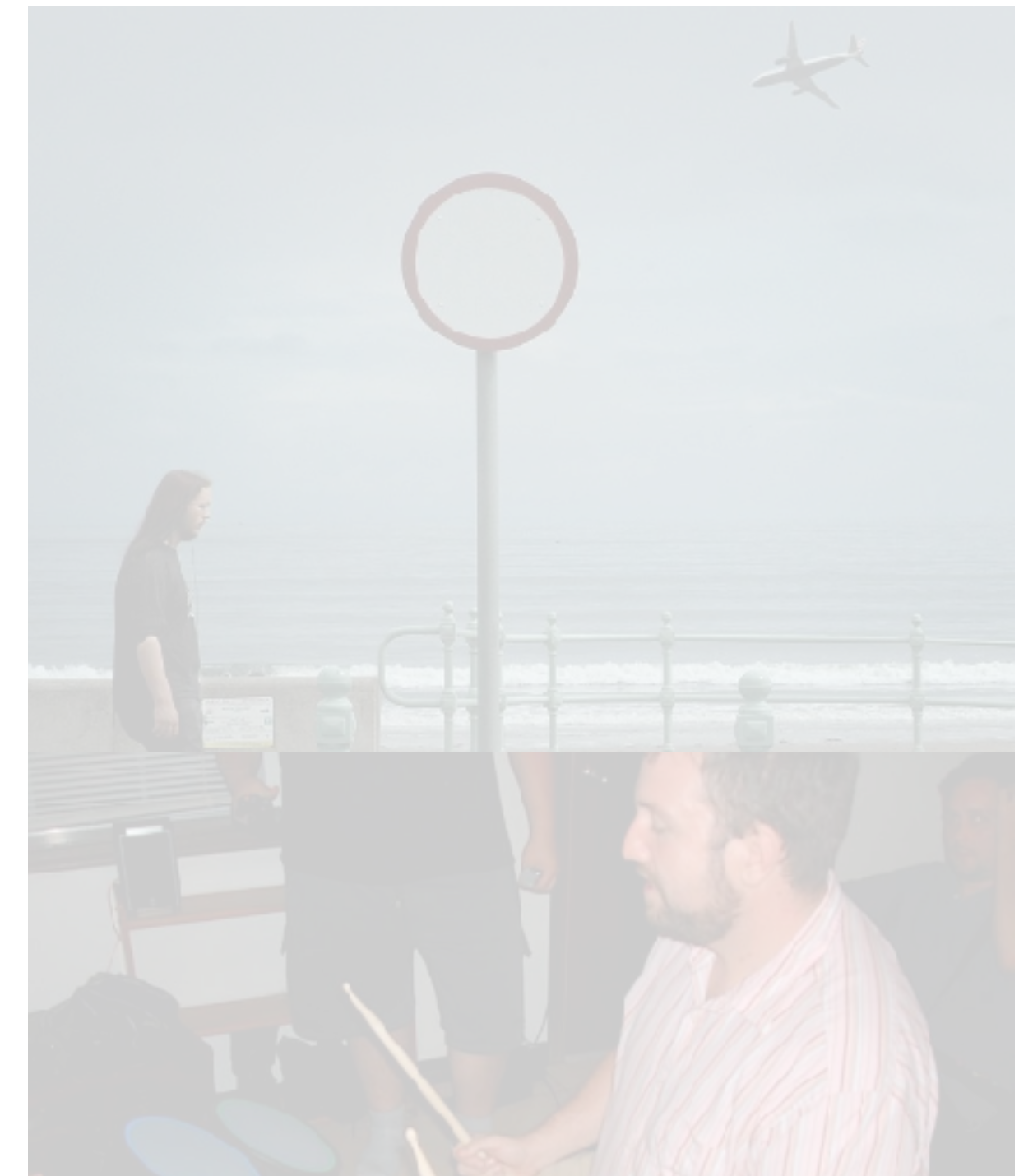
Static Webpages

Free-view
(10 sec, 30 sec)

Describe
(unlimited time)

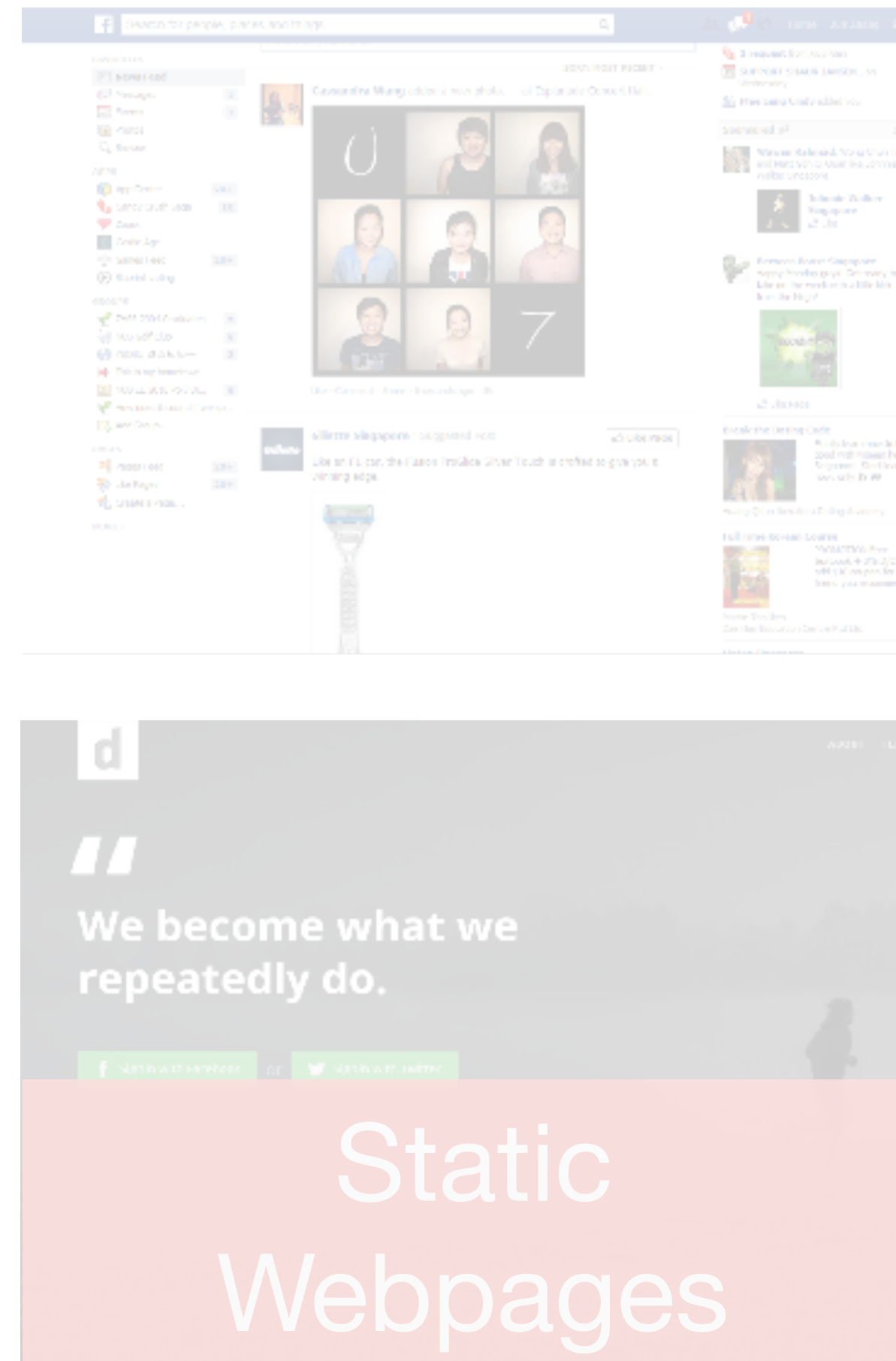
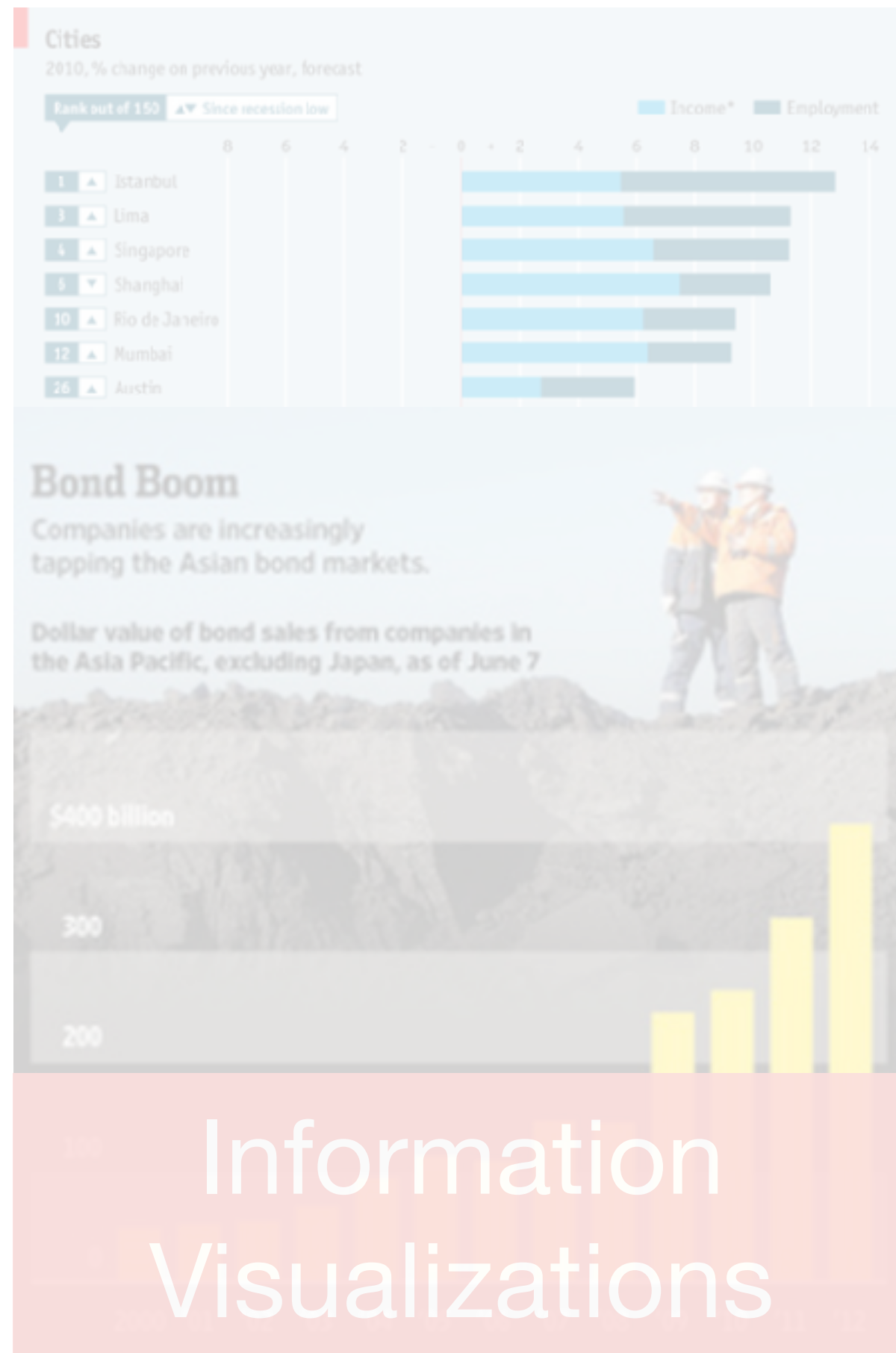
Bubble radius
(30,50,60)

vs Fixations



Natural Scene Images

Evaluation Configuration



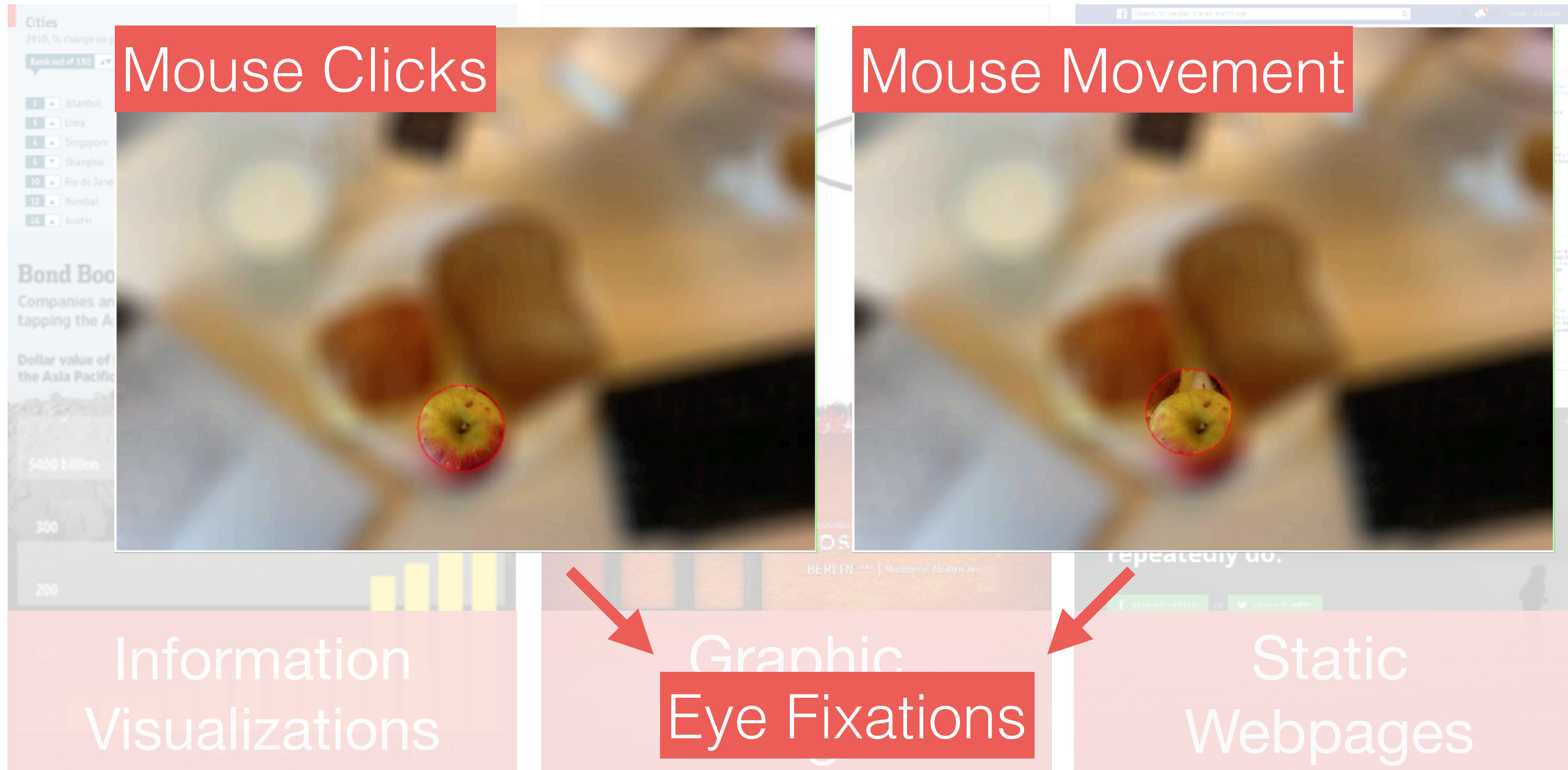
Natural Scene Images

Free-view

1. clicks (10 sec)
2. movements (5 sec)

vs Fixations

Evaluation Configuration



Natural Scene Images

Free-view

1. *clicks*
(10 sec)

2. *movements*
(5 sec)

vs Fixations

Evaluation Configuration

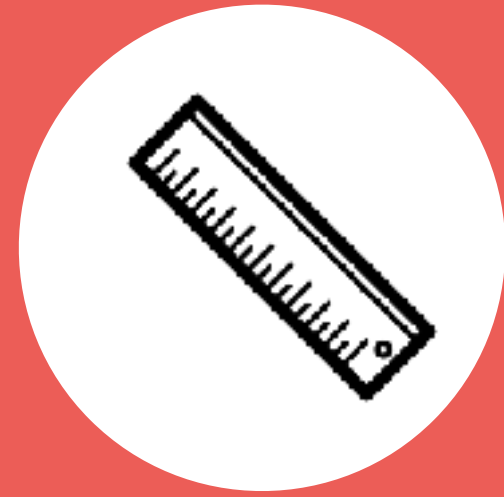
Information Visualizations	Graphic Designs	Static Webpages	Natural Scene Images
Describe (unlimited time) Bubble radius (16, 24, 32, 40)	Free-view (10 sec)	Free-view (10 sec, 30 sec) Describe (unlimited time) Bubble radius (30, 50, 60)	Free-view 1. clicks (10 sec) 2. movements (5 sec)
vs Fixations	vs Annotations	vs Fixations	vs Fixations

MASSVIS
[Borkin et al. 2016]

GDI
[O'Donovan et al. 2014]

FIWI
[Shen and Zhao 2014]

OSIE
[Xu et al. 2014]



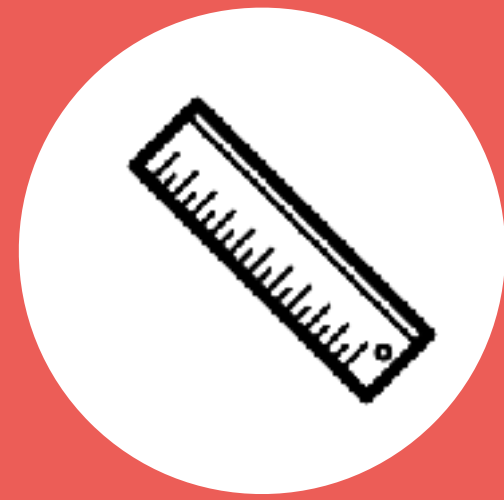
Evaluation Tools



Experimental Results



Future Applications



Evaluation Tools



Experimental Results



Future Applications



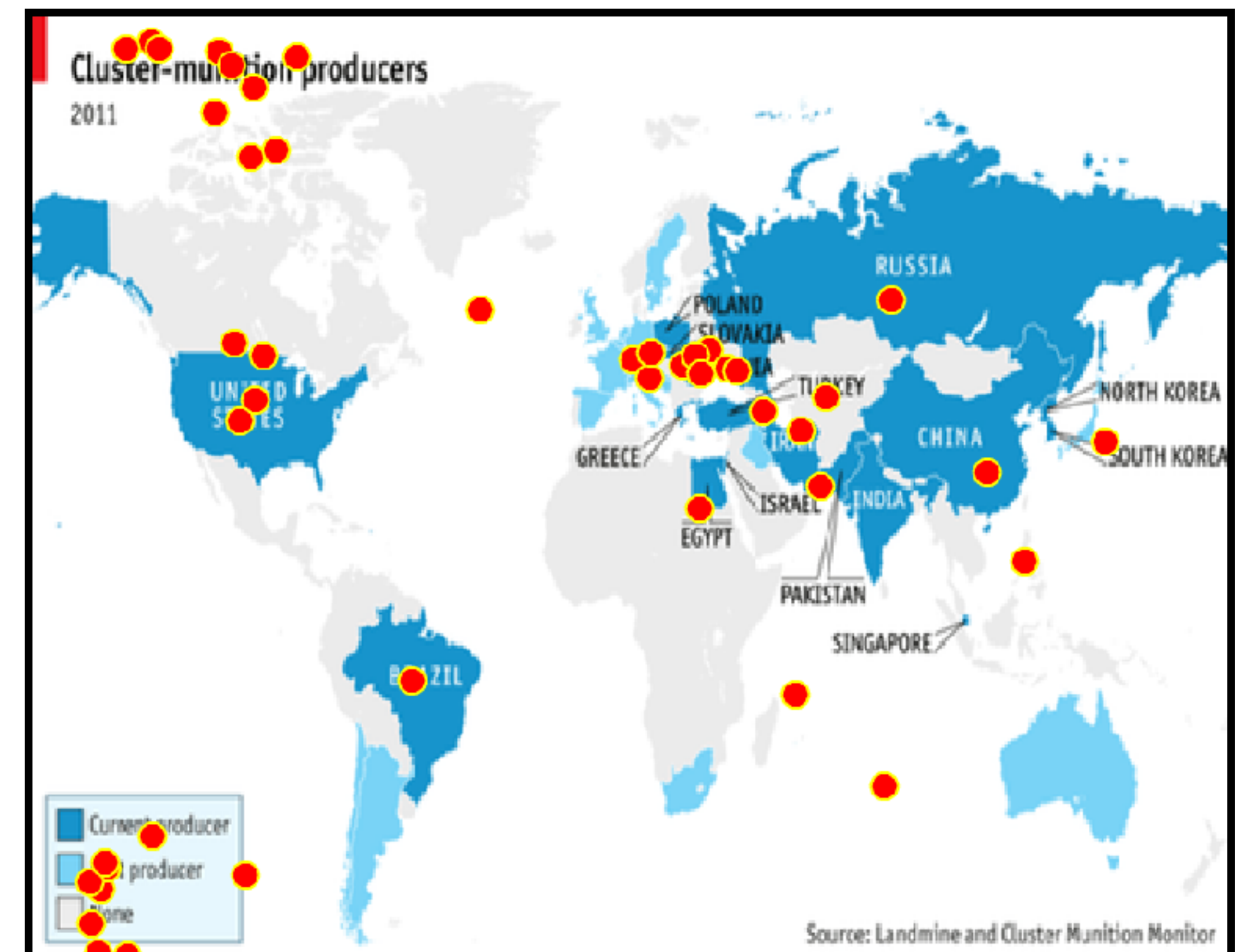
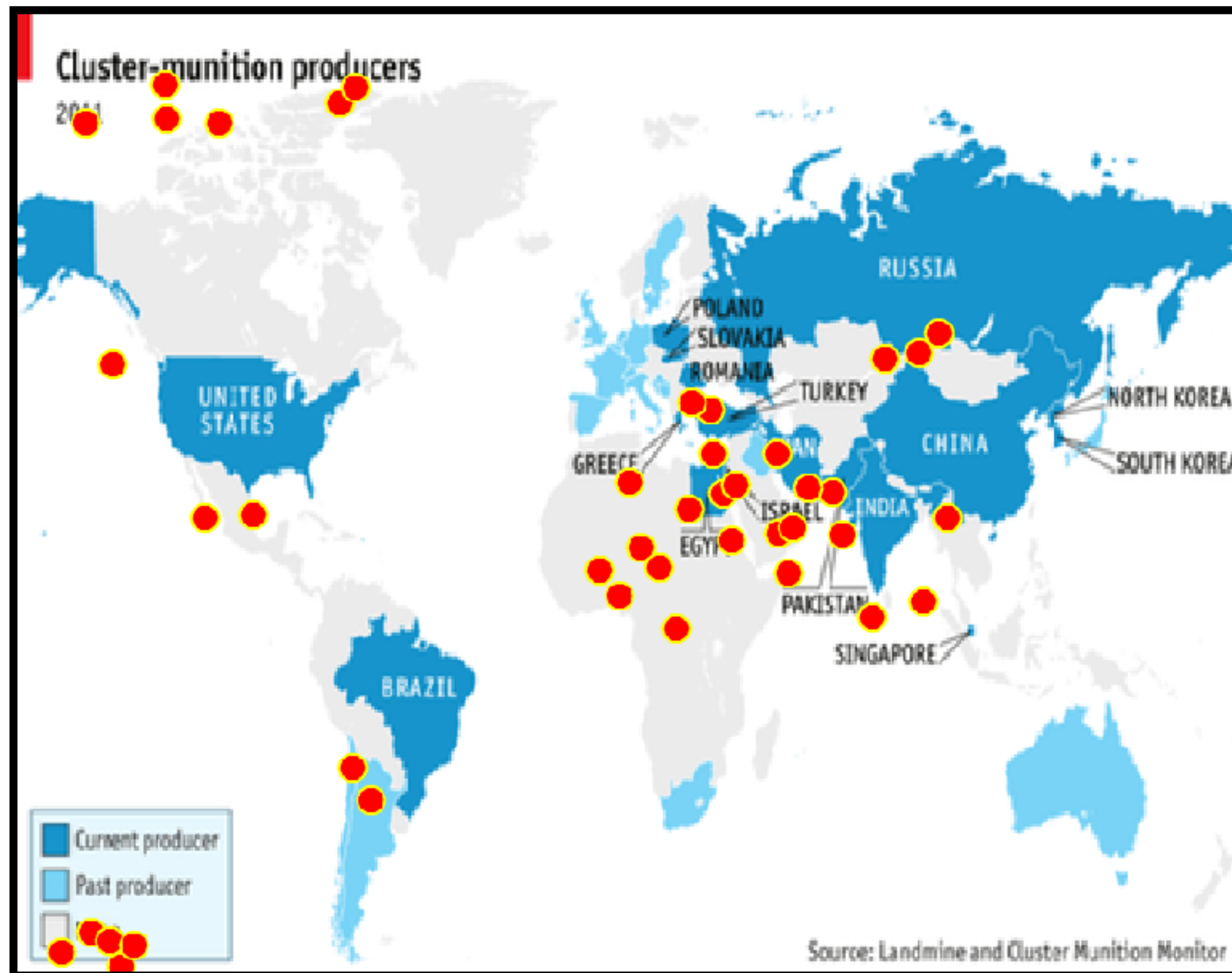
Computing CC score



Clicks



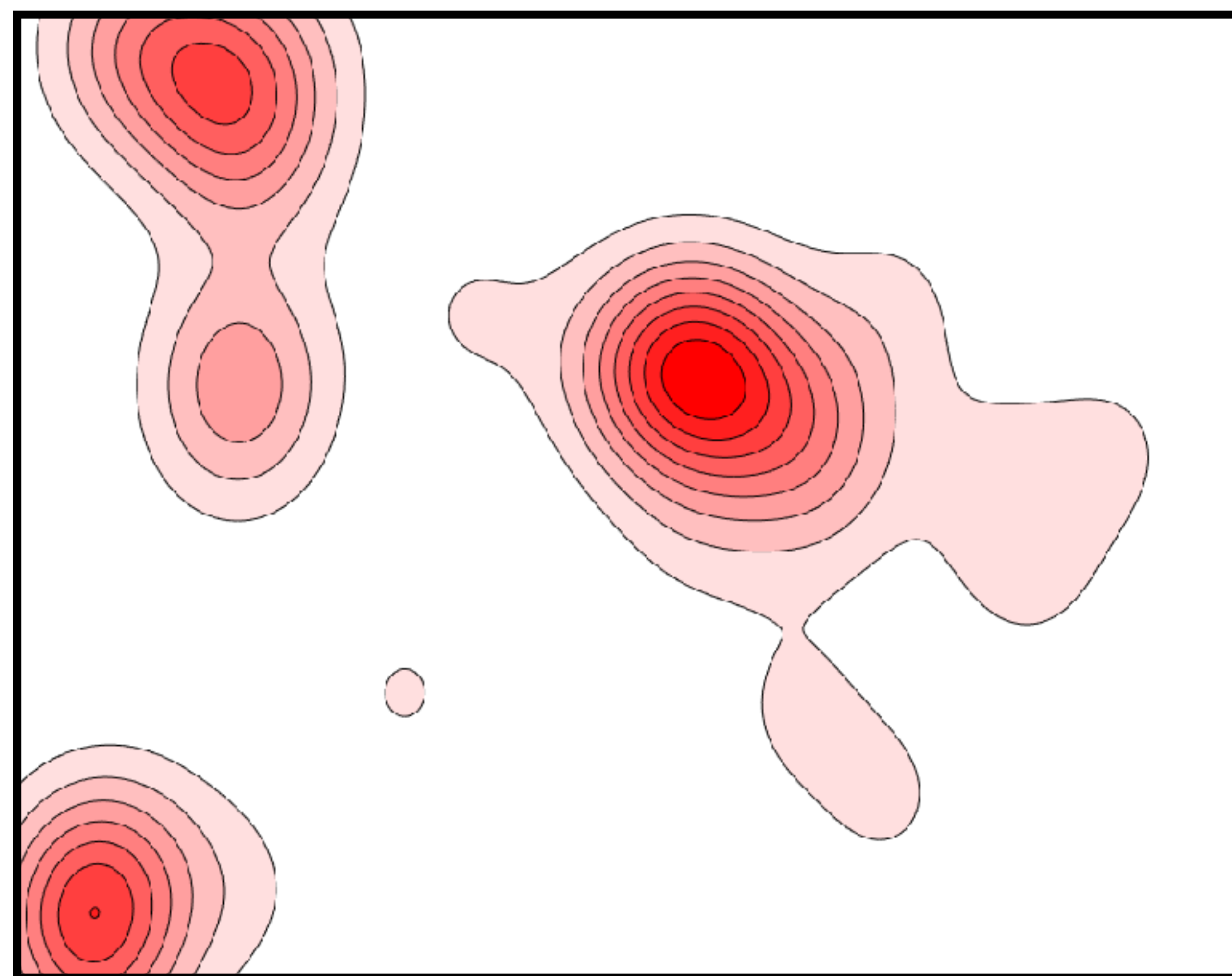
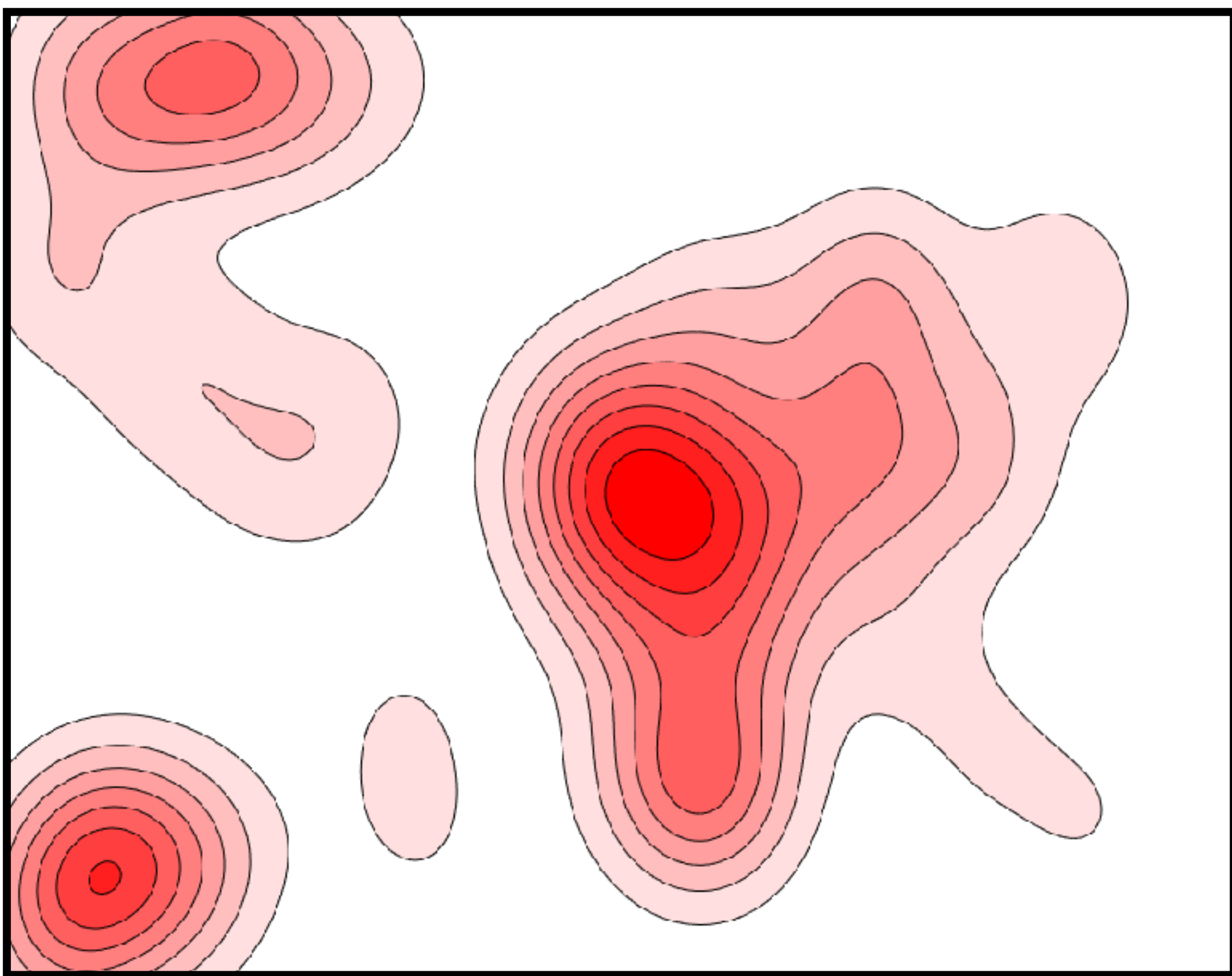
Fixations



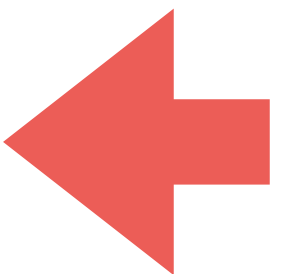
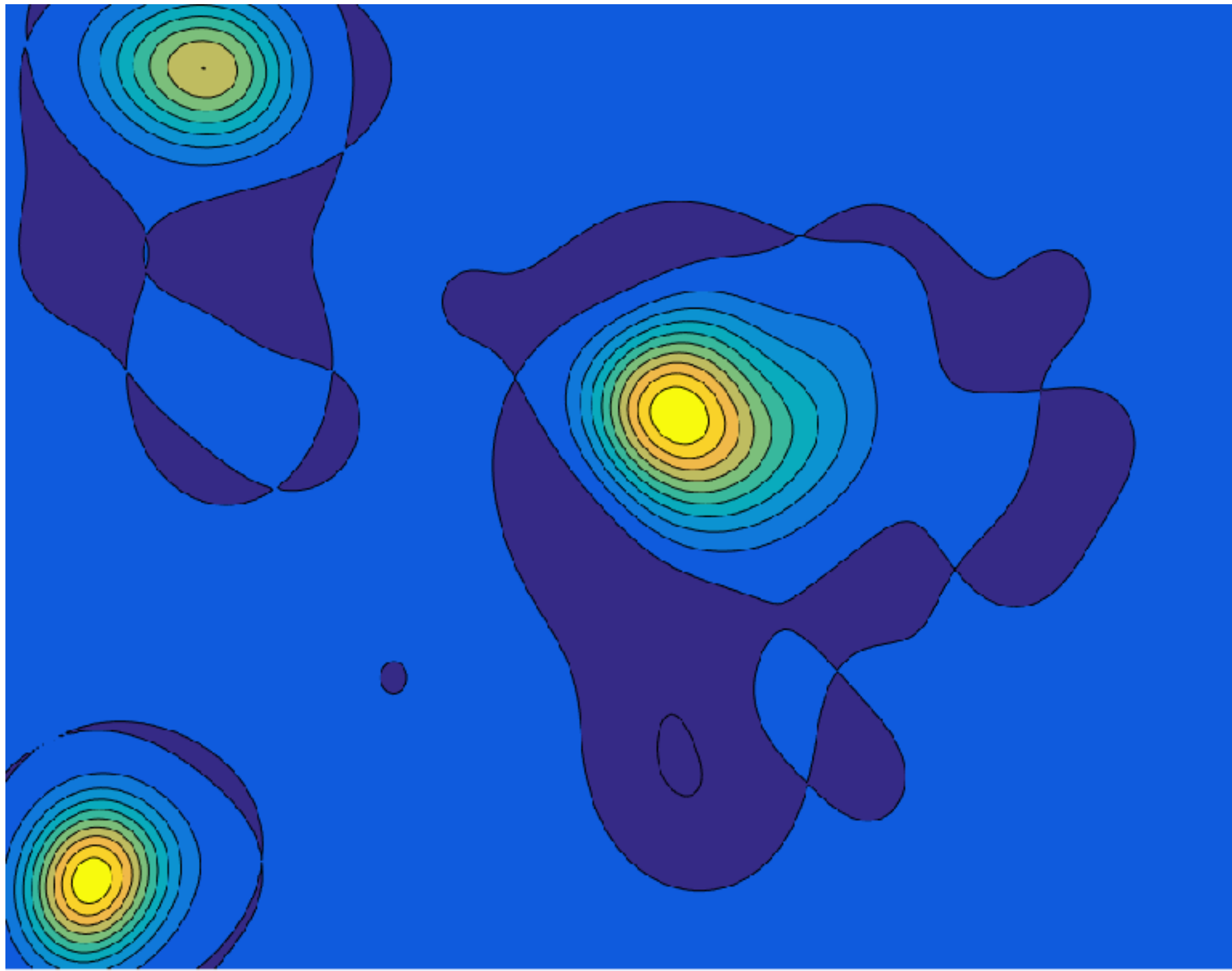
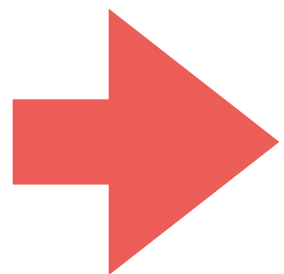
Computing CC score

 Clicks

 Fixations



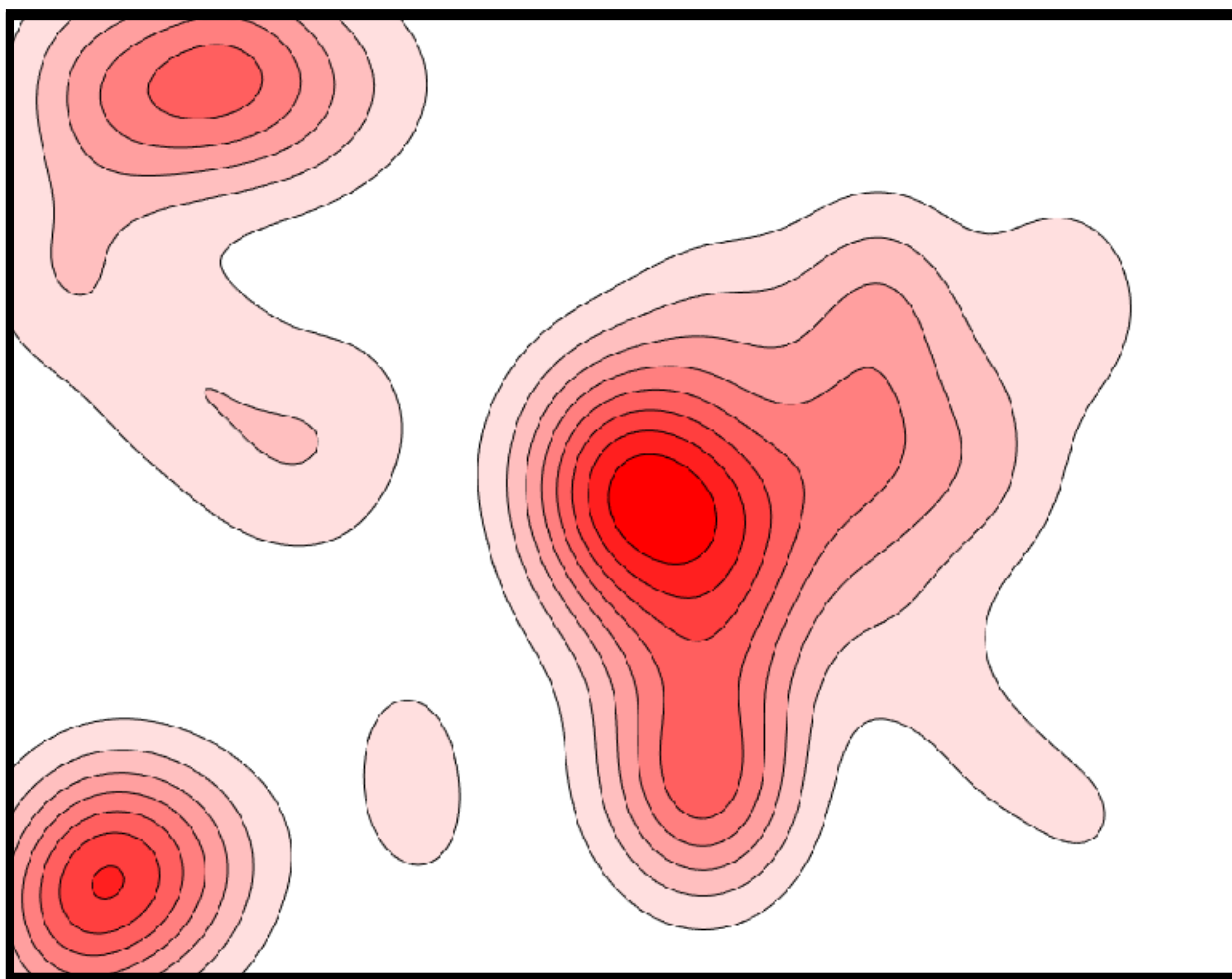
Computing CC score



Computing NSS score

 Clicks

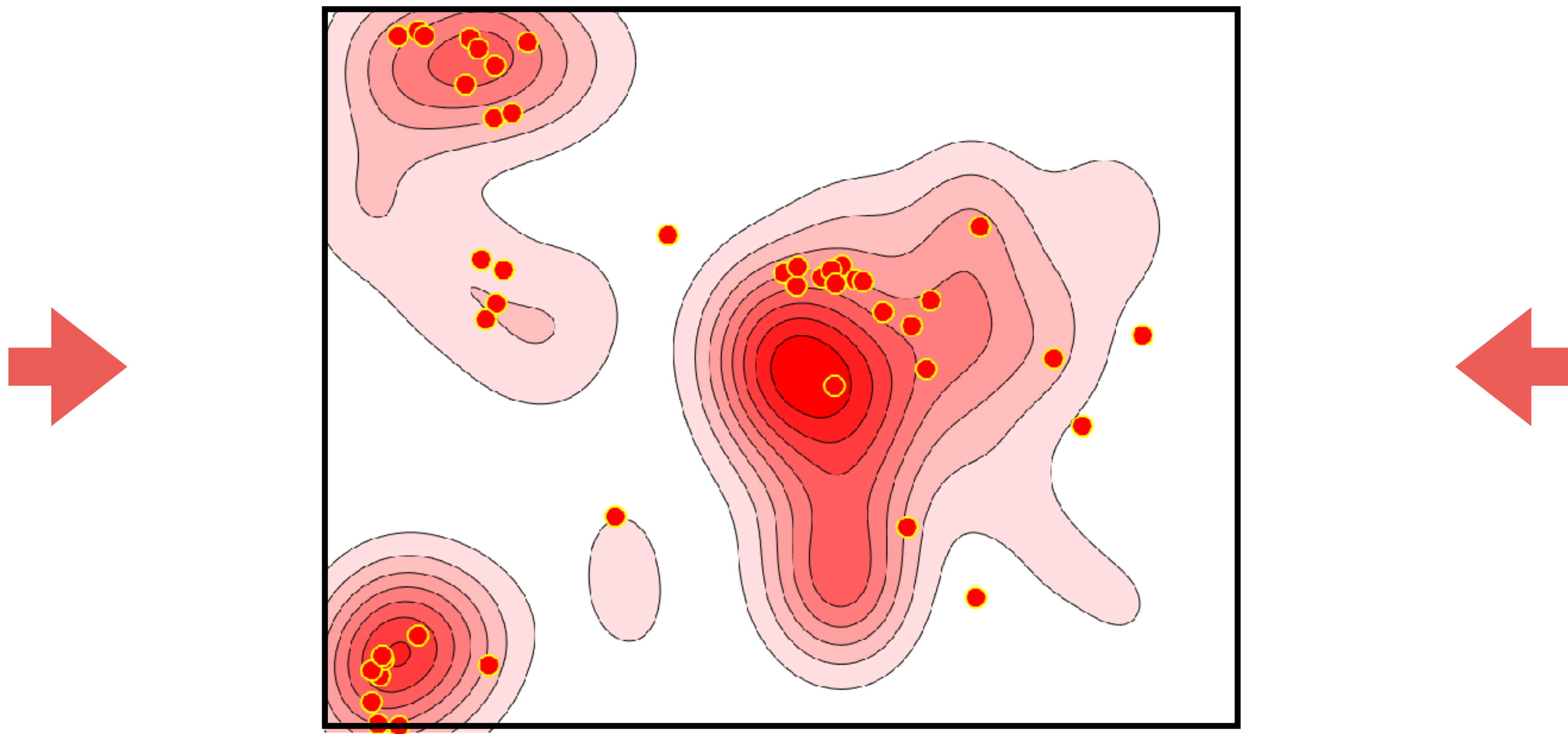
 Fixations





Computing NSS score

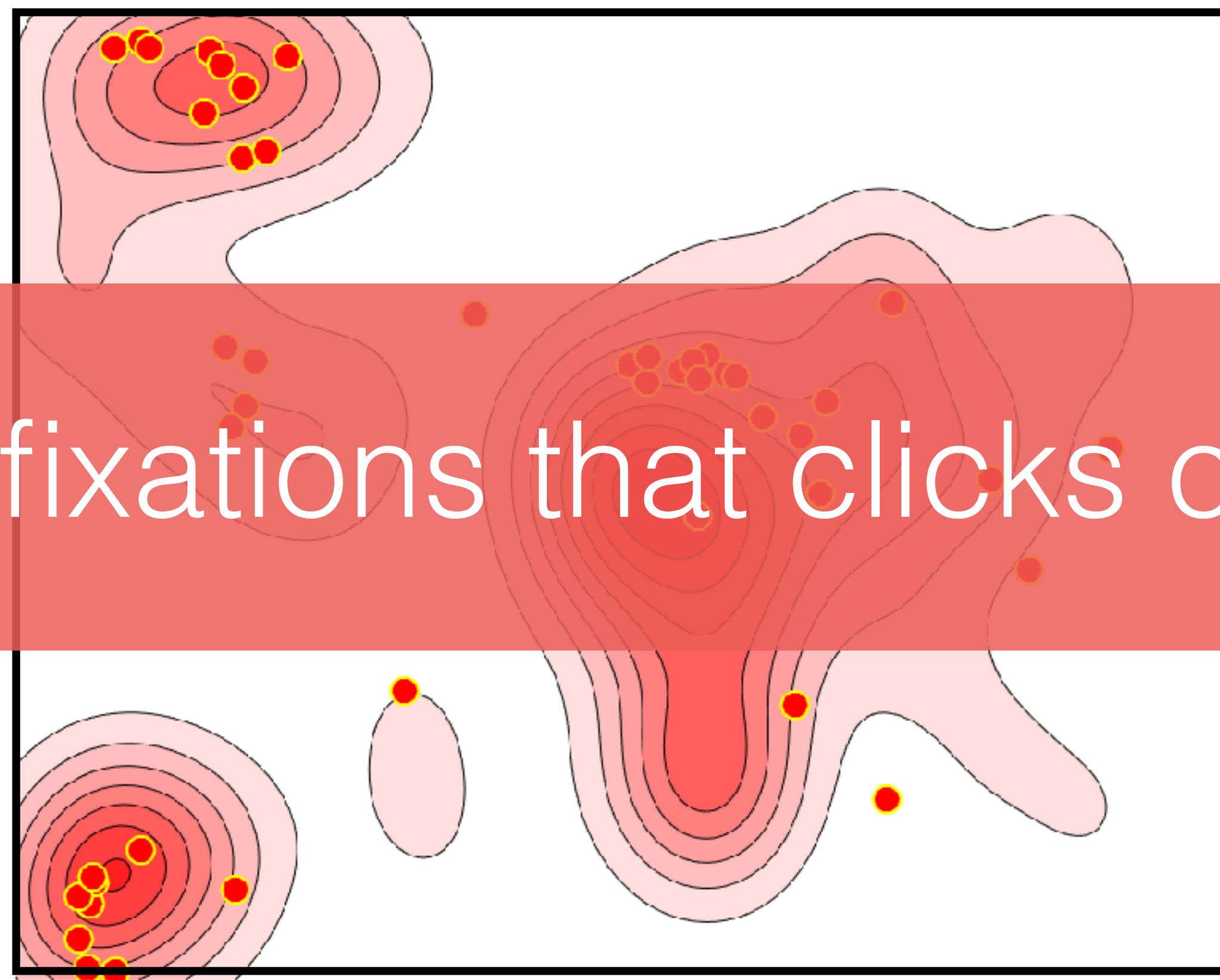
Normalized by eye fixation consistency





Computing NSS score

Normalized by eye fixation consistency



Report % of fixations that clicks can explain



Evaluation Tools



Experimental Results



Future Applications



Take-away #1:

Clicks are more effective than mouse movements for measuring observer behavior.

Clicks vs Movements



Movements

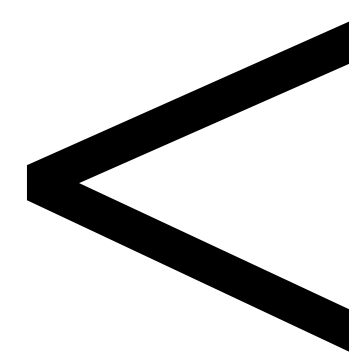


Clicks

Clicks are conscious decisions of importance



Movements



Clicks





Clicks are conscious decisions of importance



Clicks are a better approximation to eye fixations

Movements



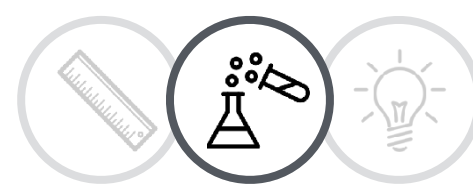
Intentionality

Clicks



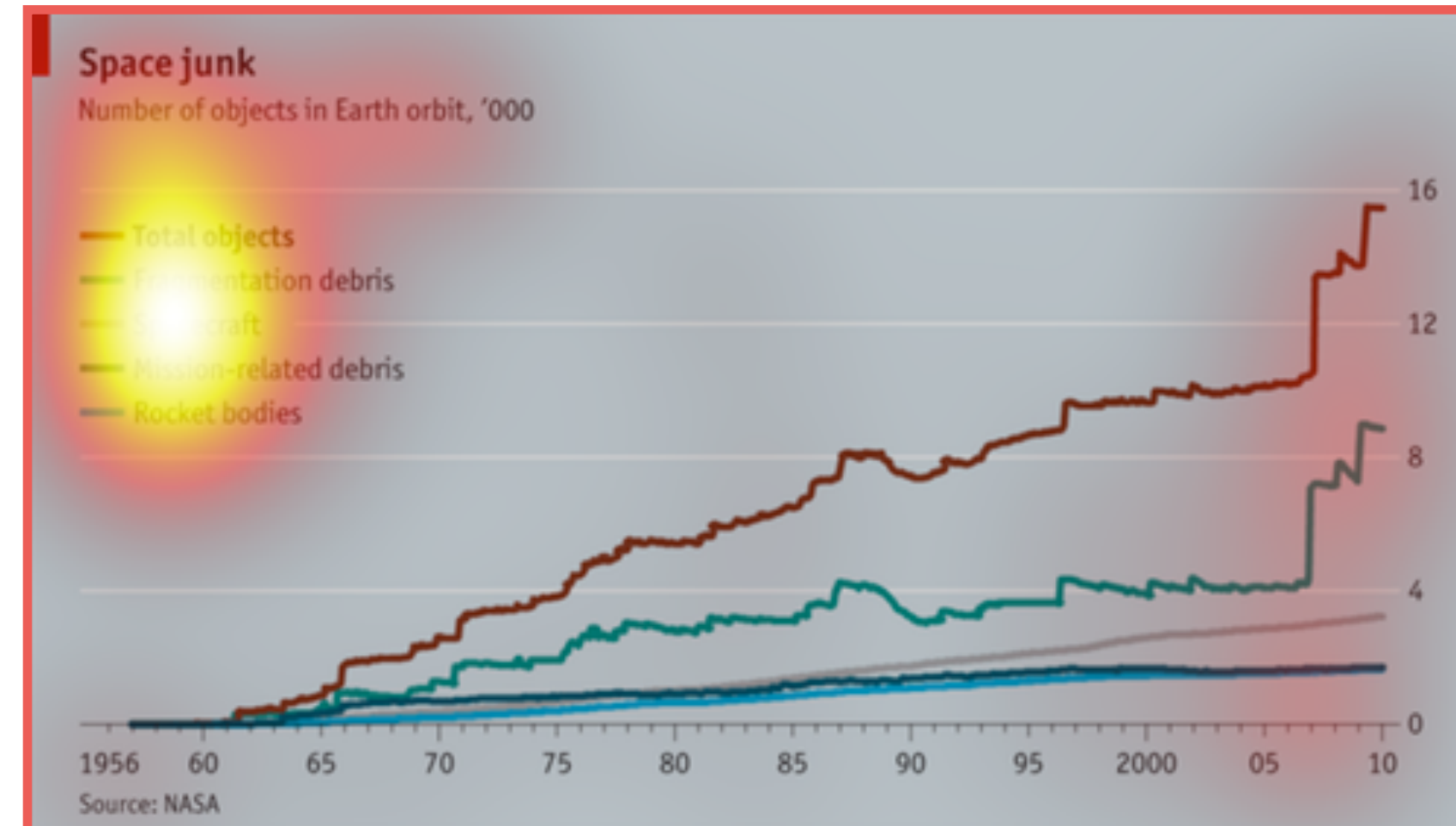
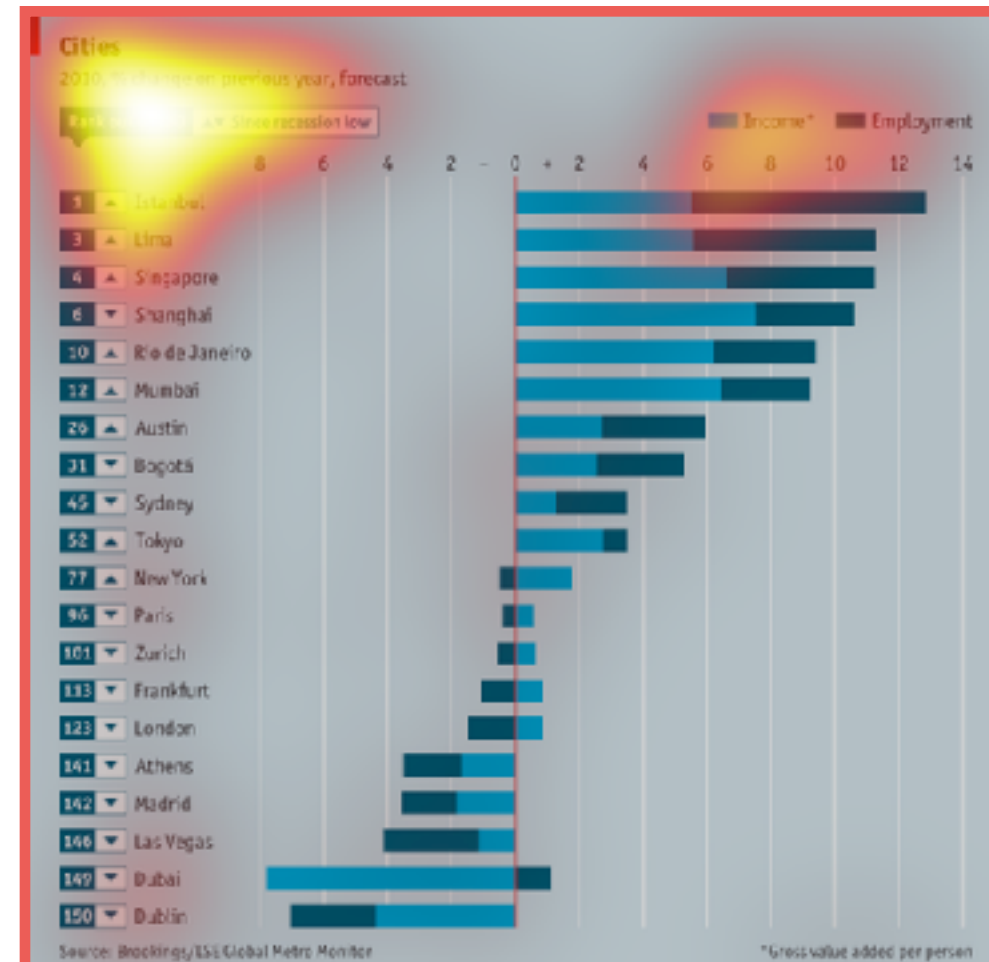
Take-away #2:

Clicks are predictive of eye fixations across a variety of image types and tasks.



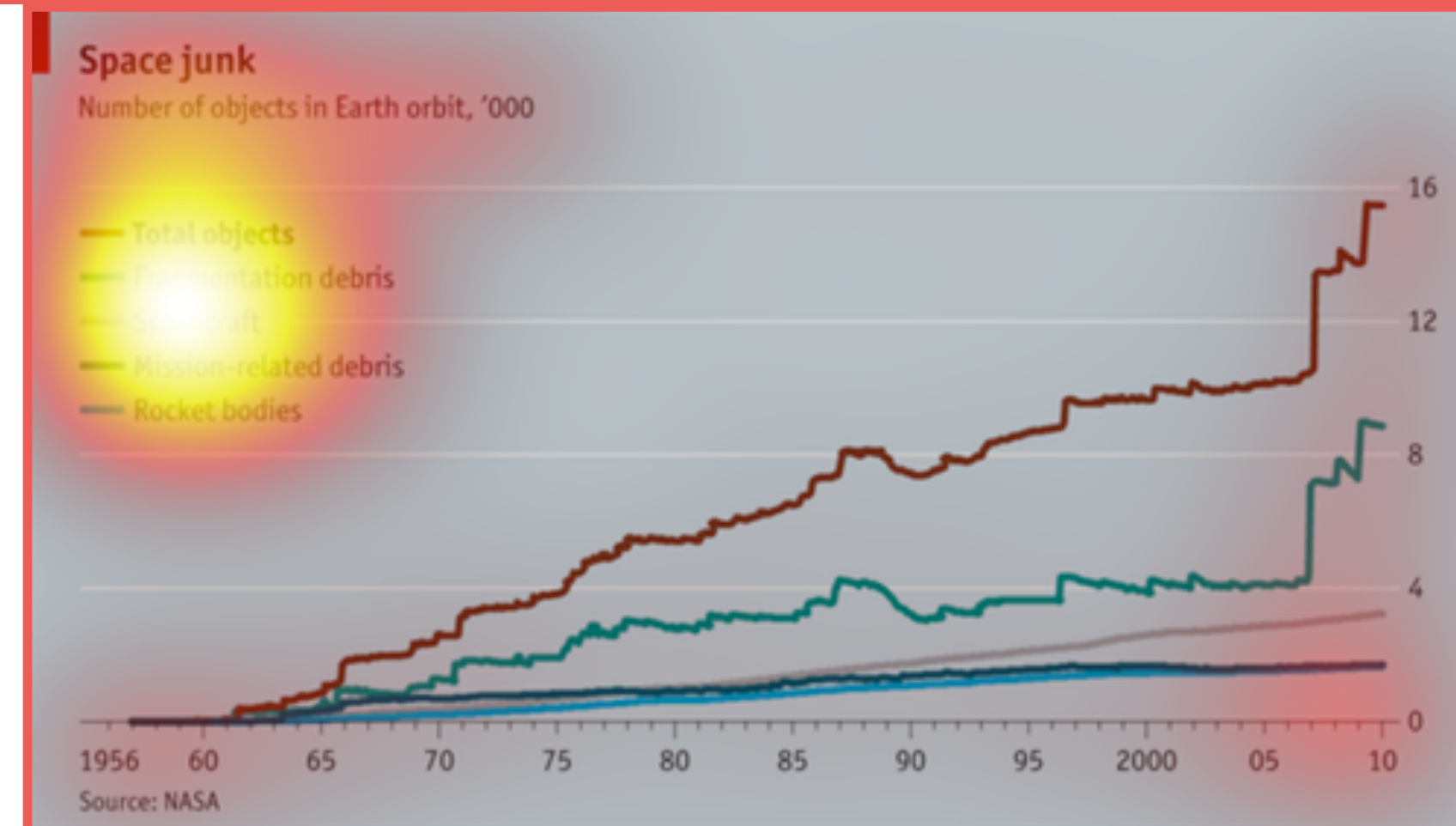
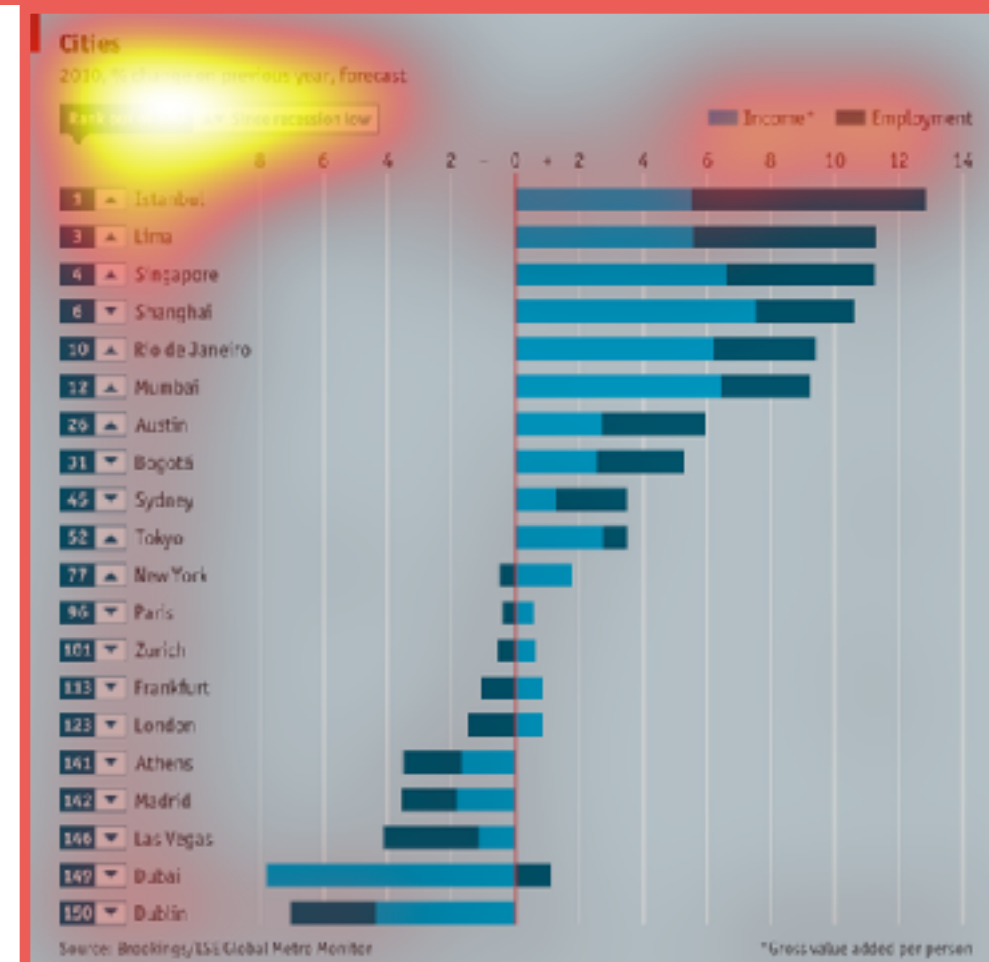
Clicks predict fixations on visualizations

Fixations



clicks of 10 participants explain 90% of fixations

Clicks



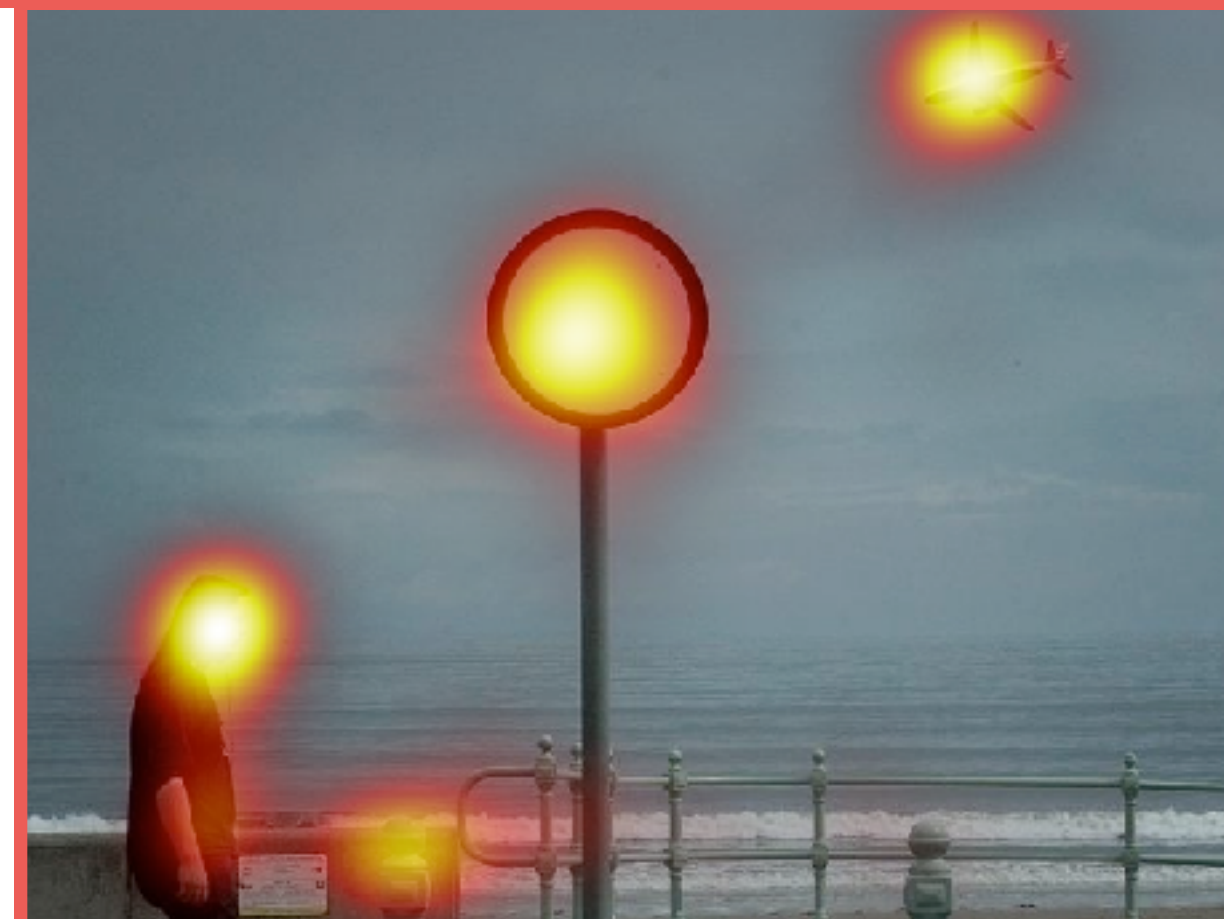
Clicks predict fixations on natural images

Fixations



clicks of 10 participants explain 78% of fixations

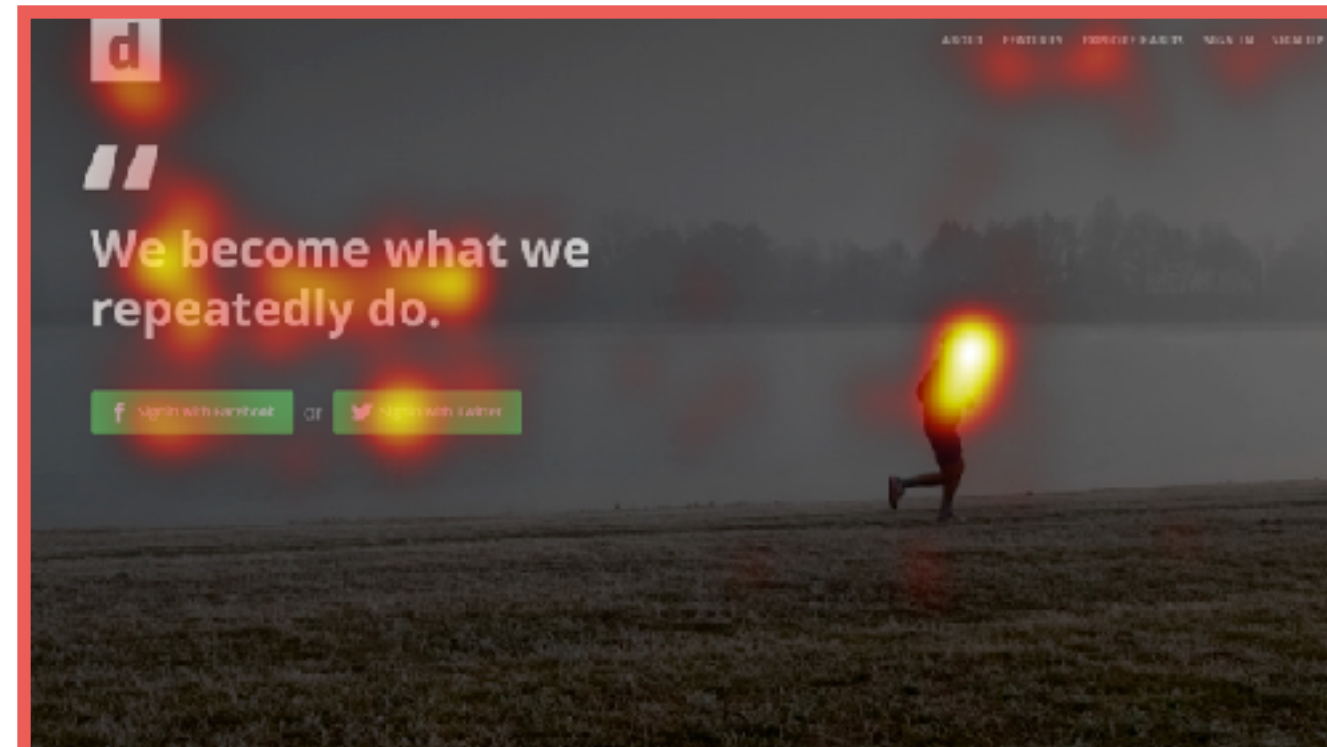
Clicks





Clicks predict fixations on webpages

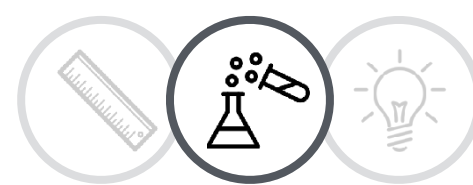
Fixations



clicks of 10 participants explain 78% of fixations

Clicks





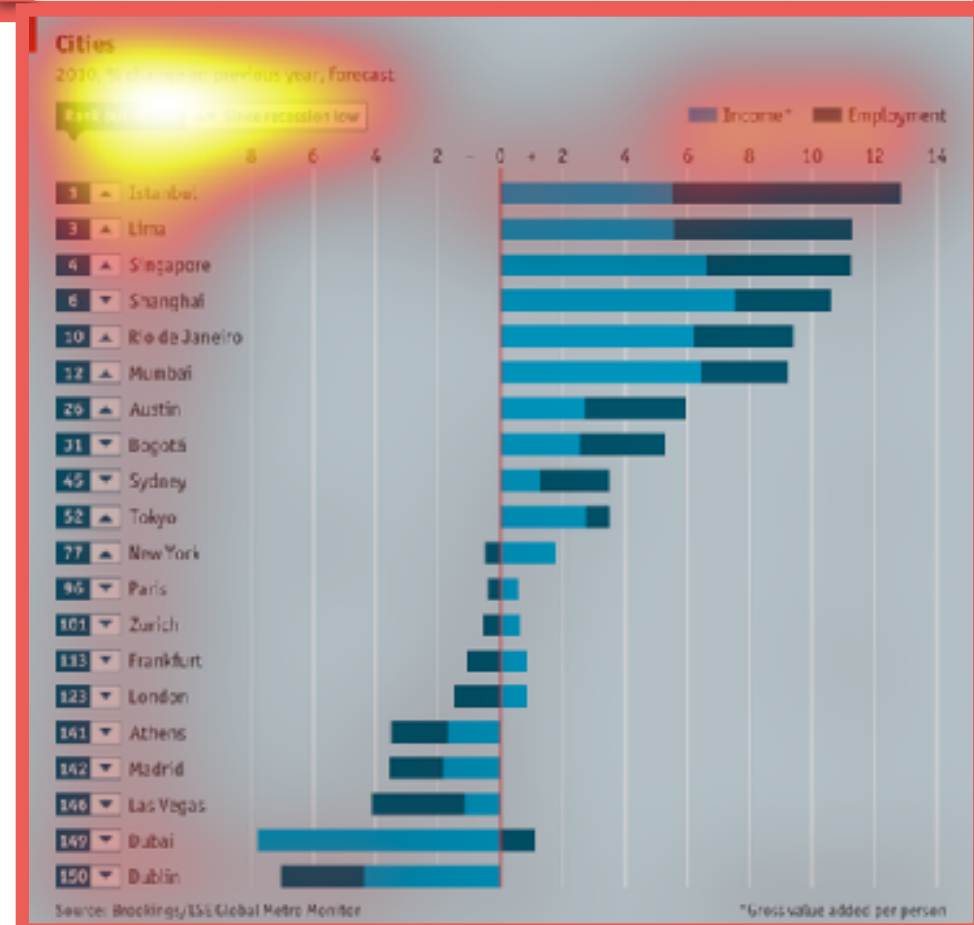
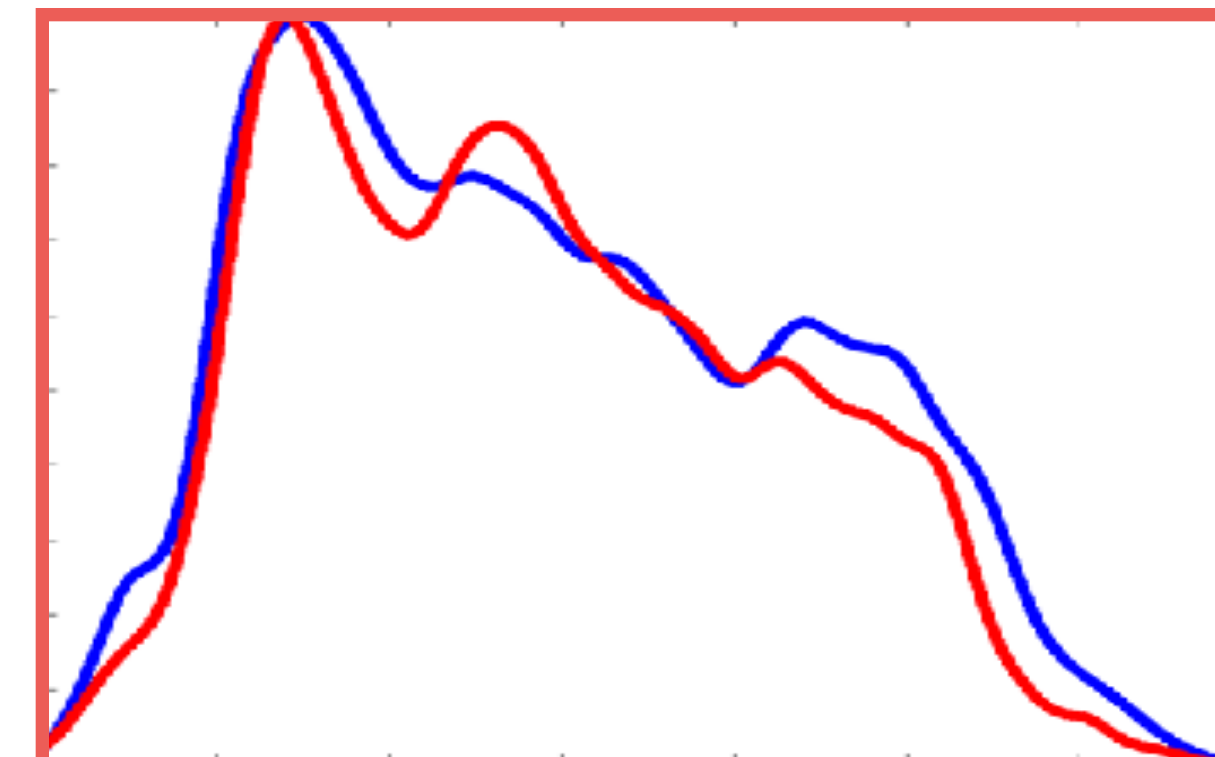
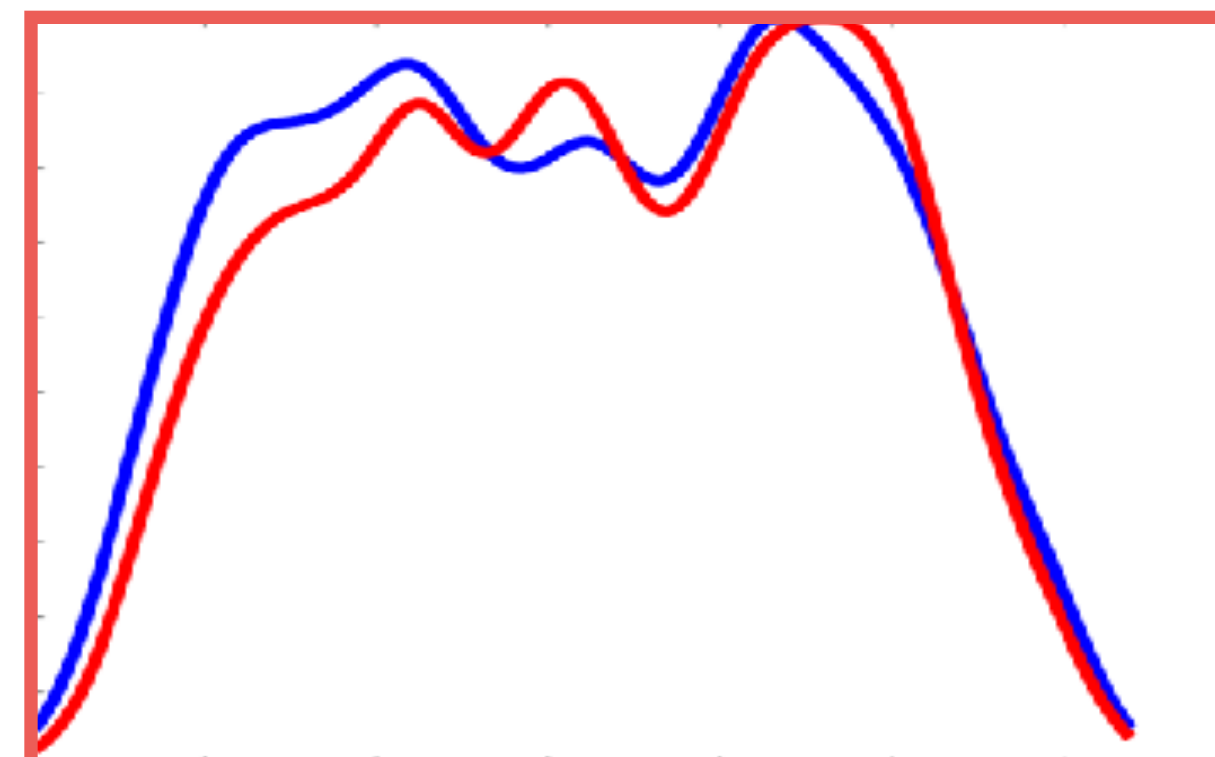
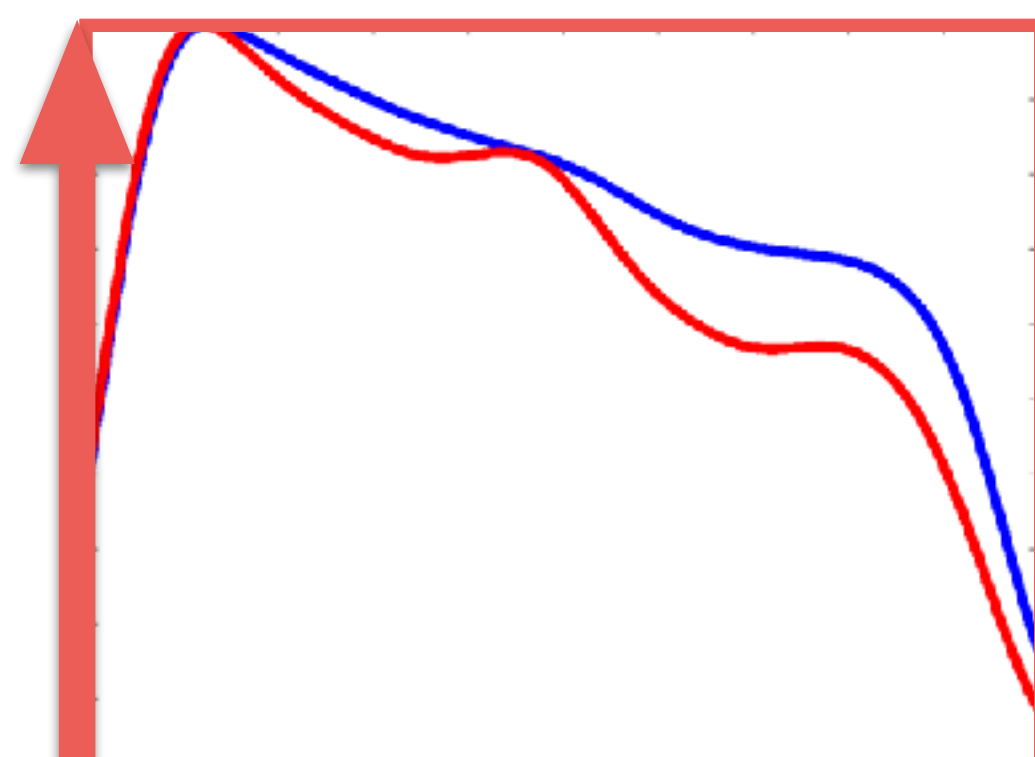
Click patterns match fixation patterns

Visualizations

Natural images

Webpages

Heatmap intensity
bubble clicks
eye fixations

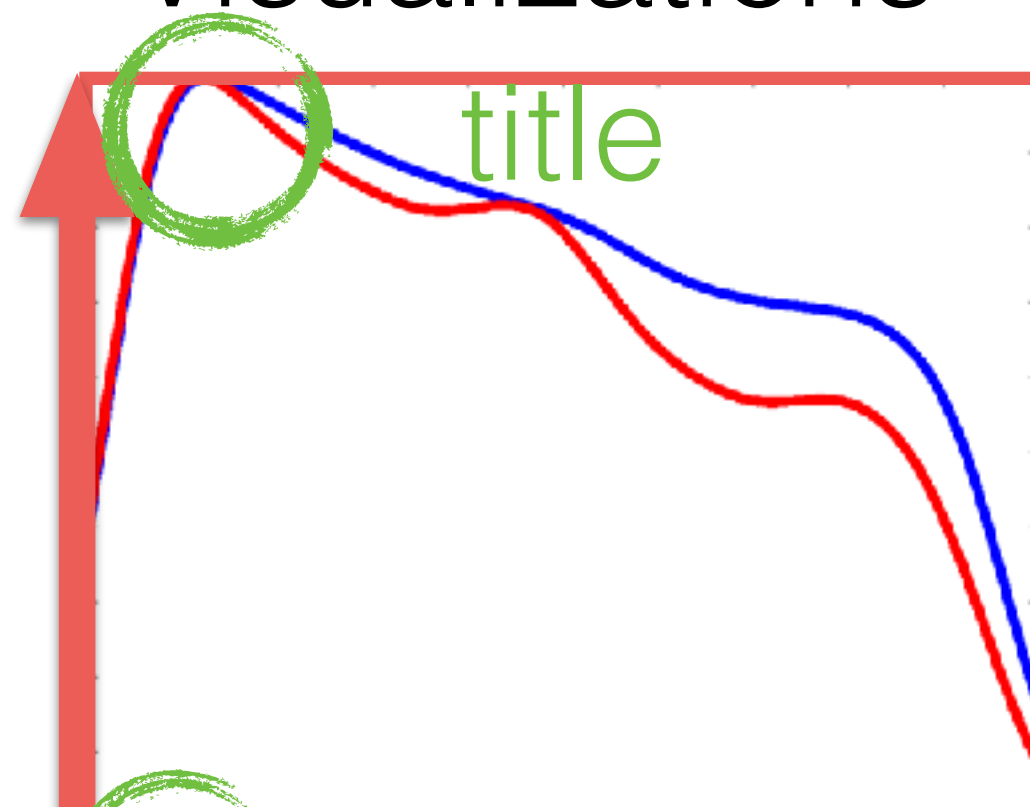




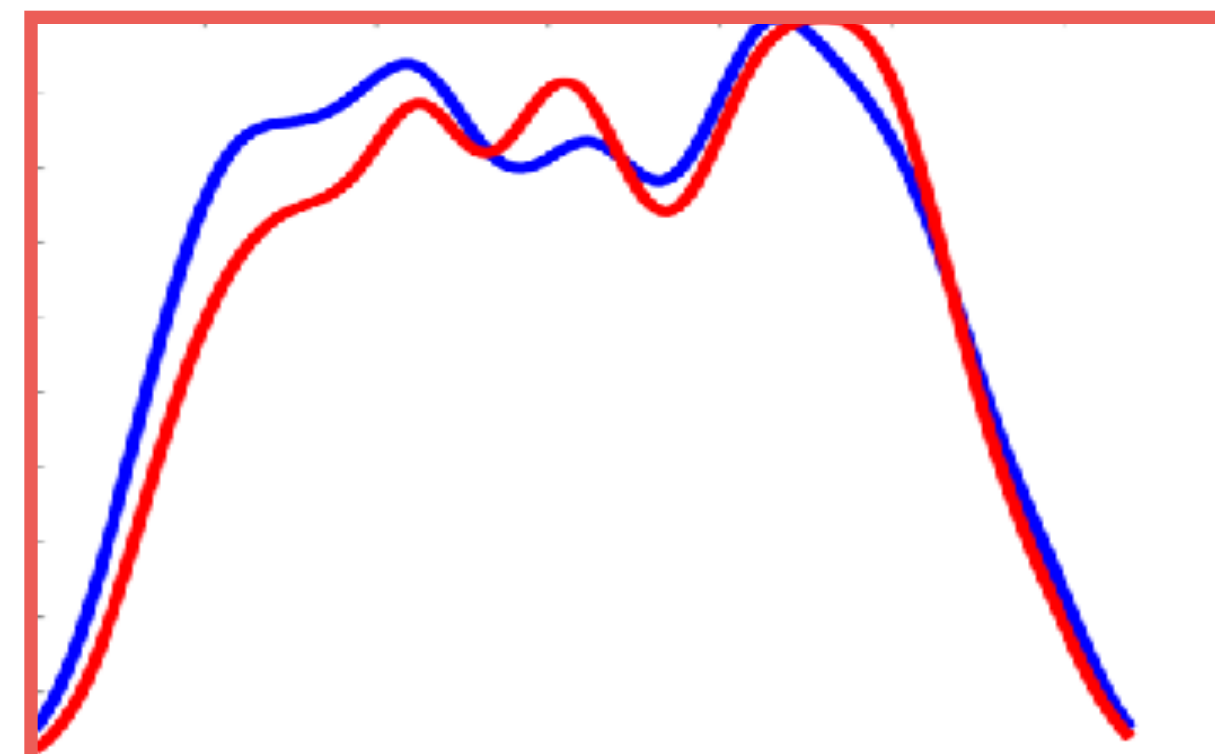
Click patterns match fixation patterns

Heatmap intensity
bubble clicks
eye fixations

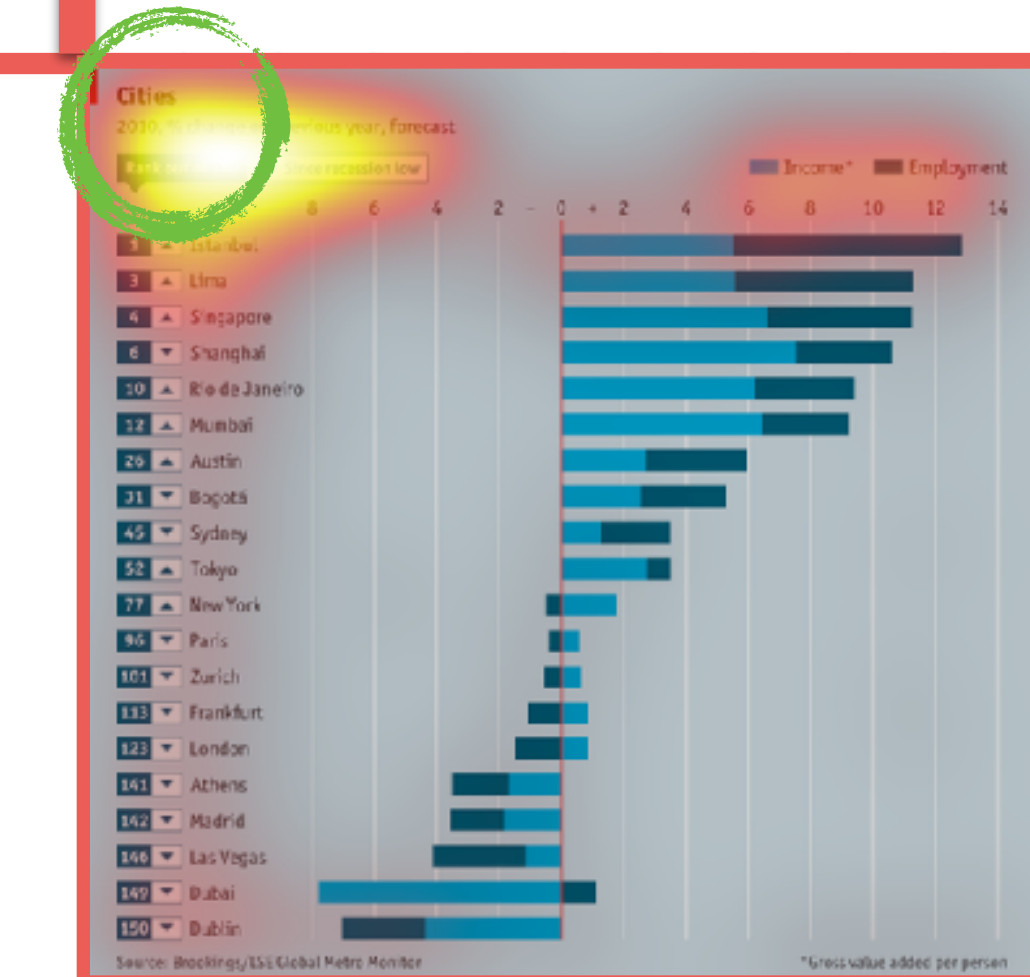
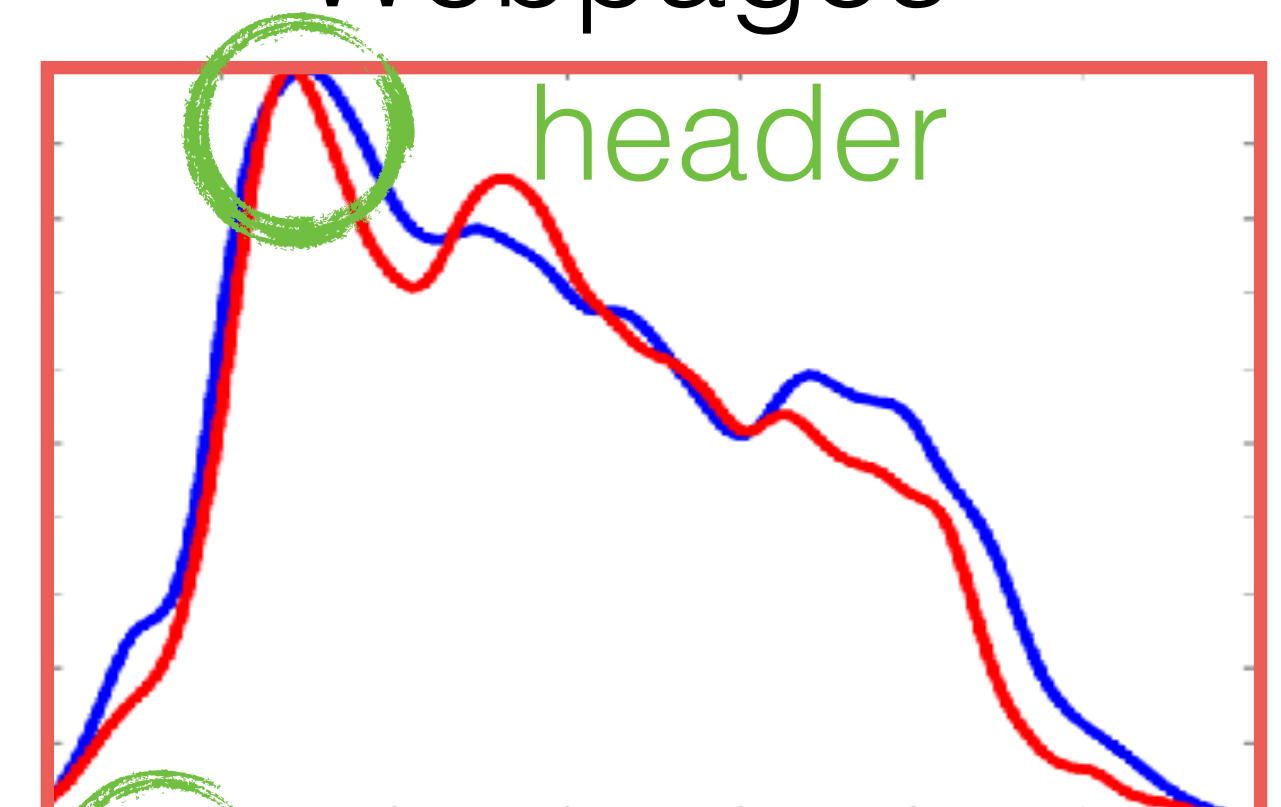
Visualizations



Natural images



Webpages





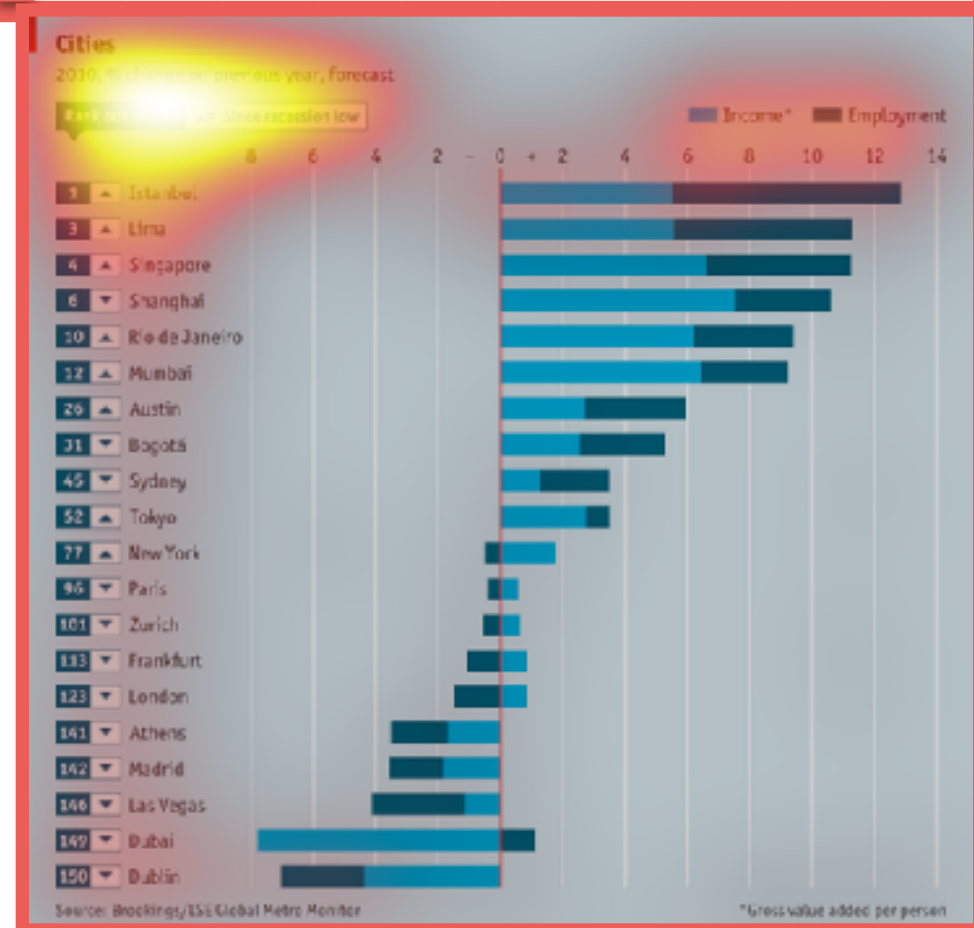
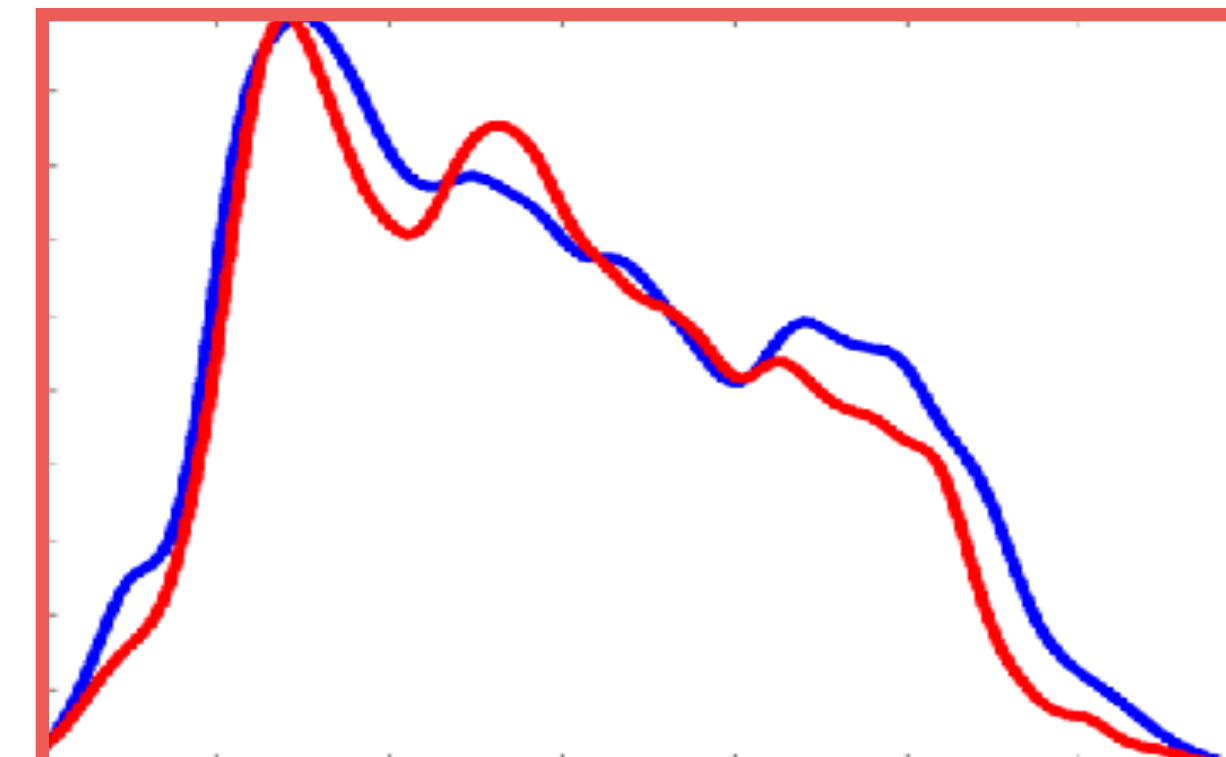
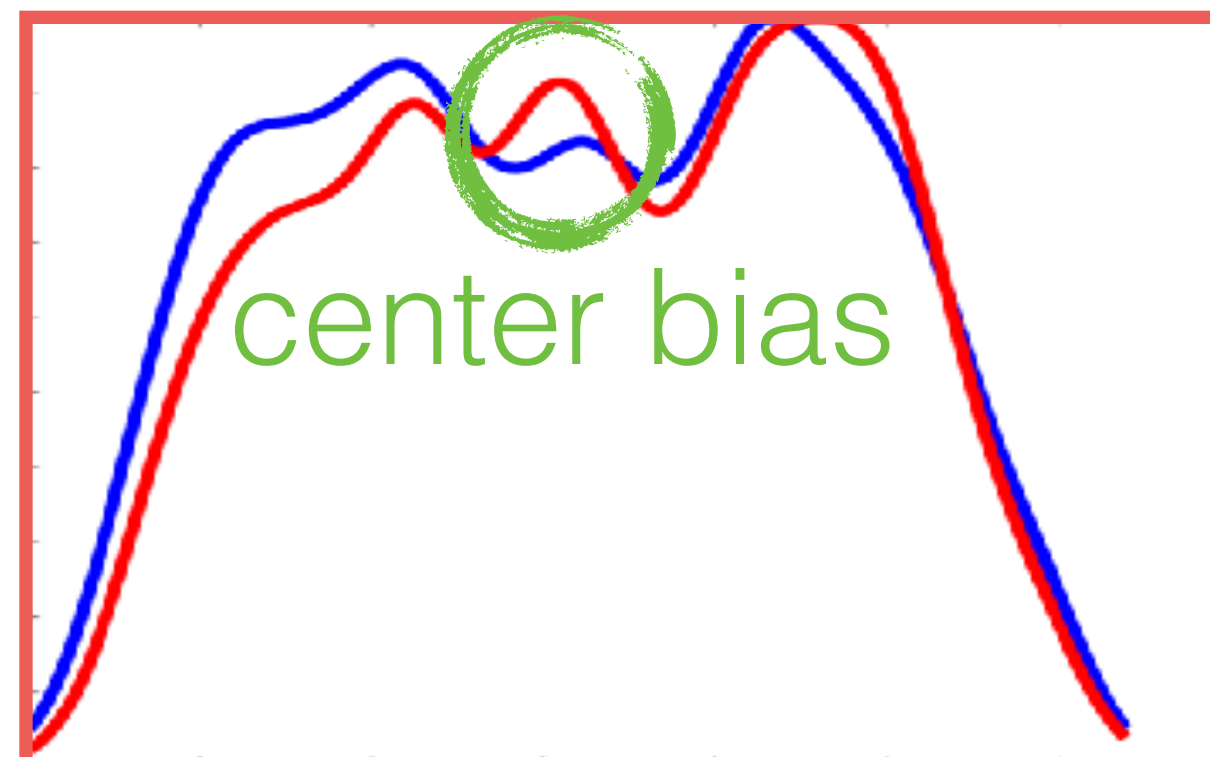
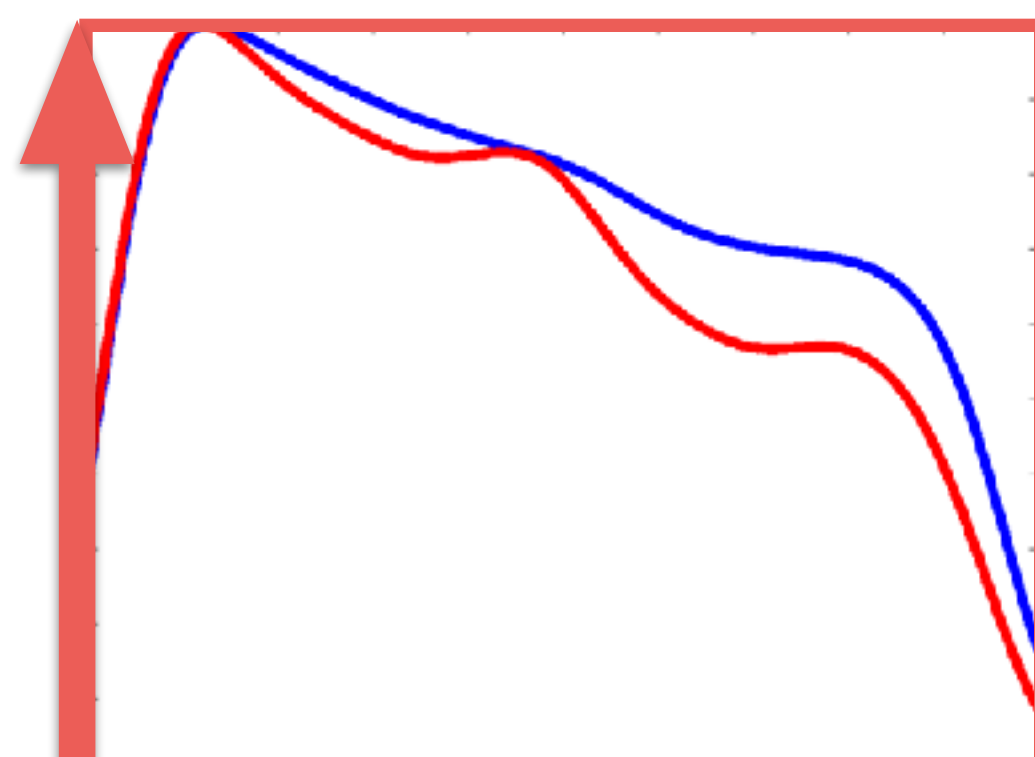
Click patterns match fixation patterns

Visualizations

Natural images

Webpages

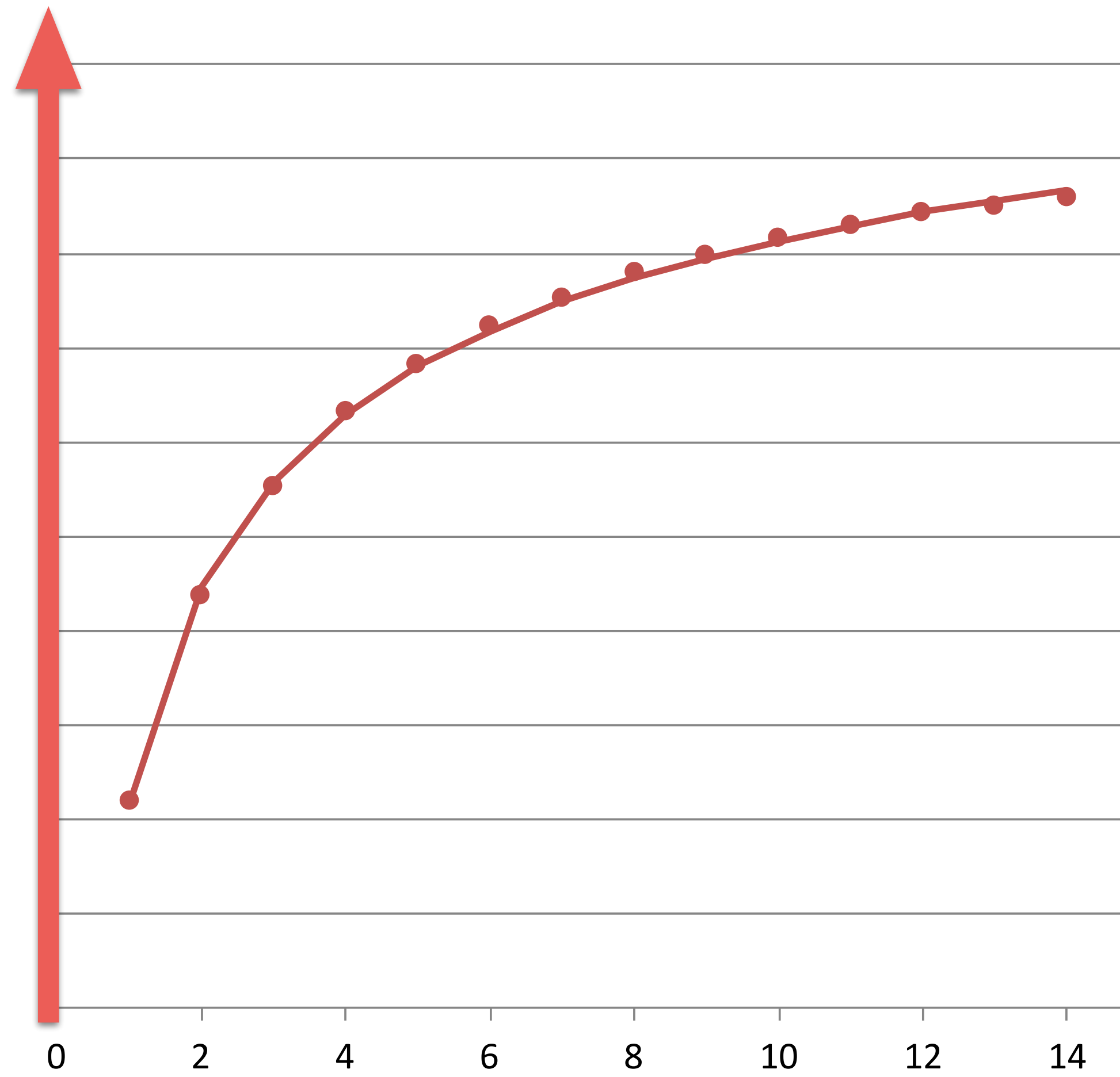
Heatmap intensity
bubble clicks
eye fixations





More involved tasks lead to better clicks

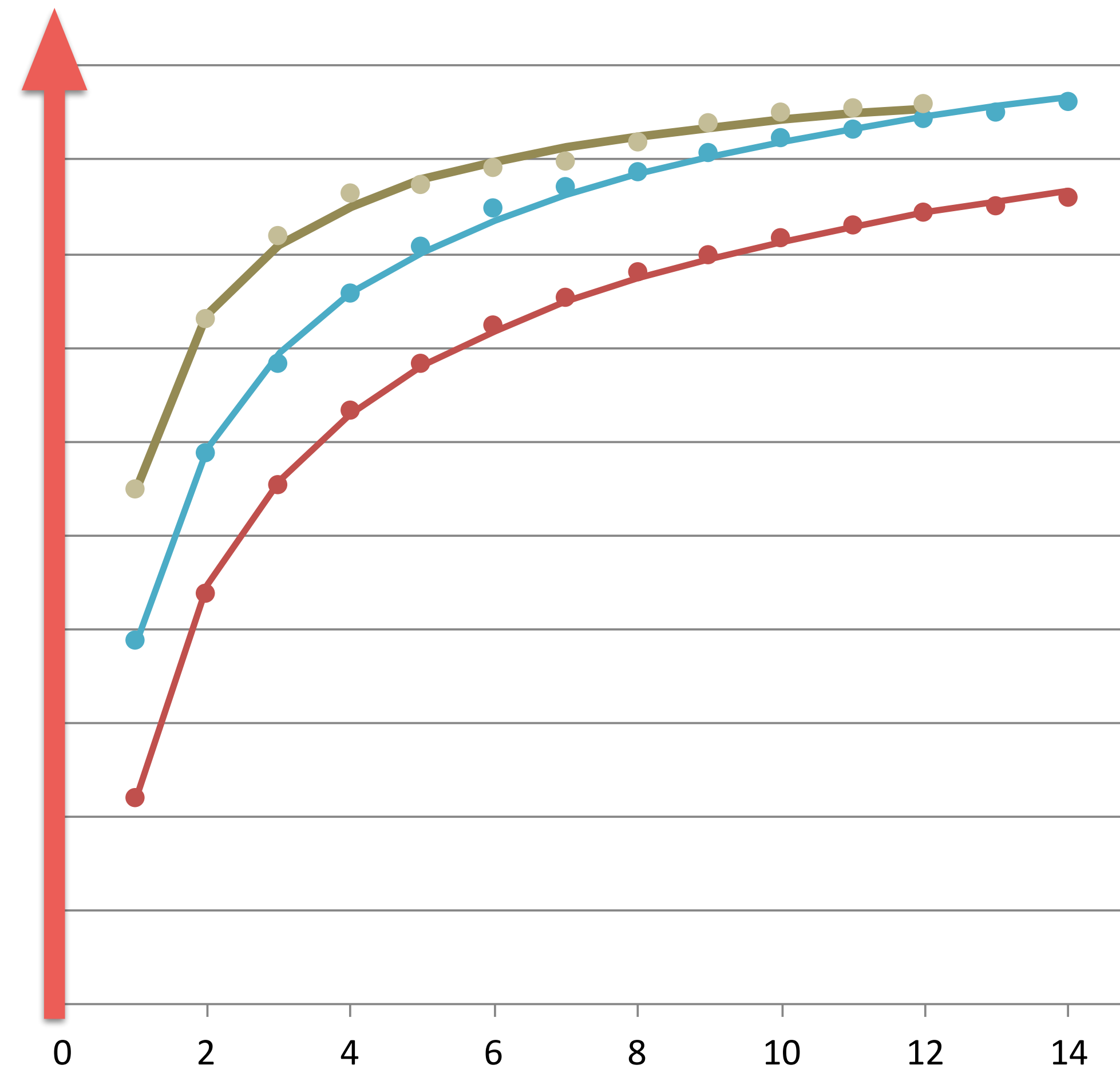
Similarity between clicks and fixations



BubbleView participants

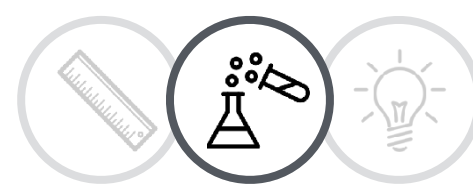
More involved tasks lead to better clicks

Similarity between clicks and fixations



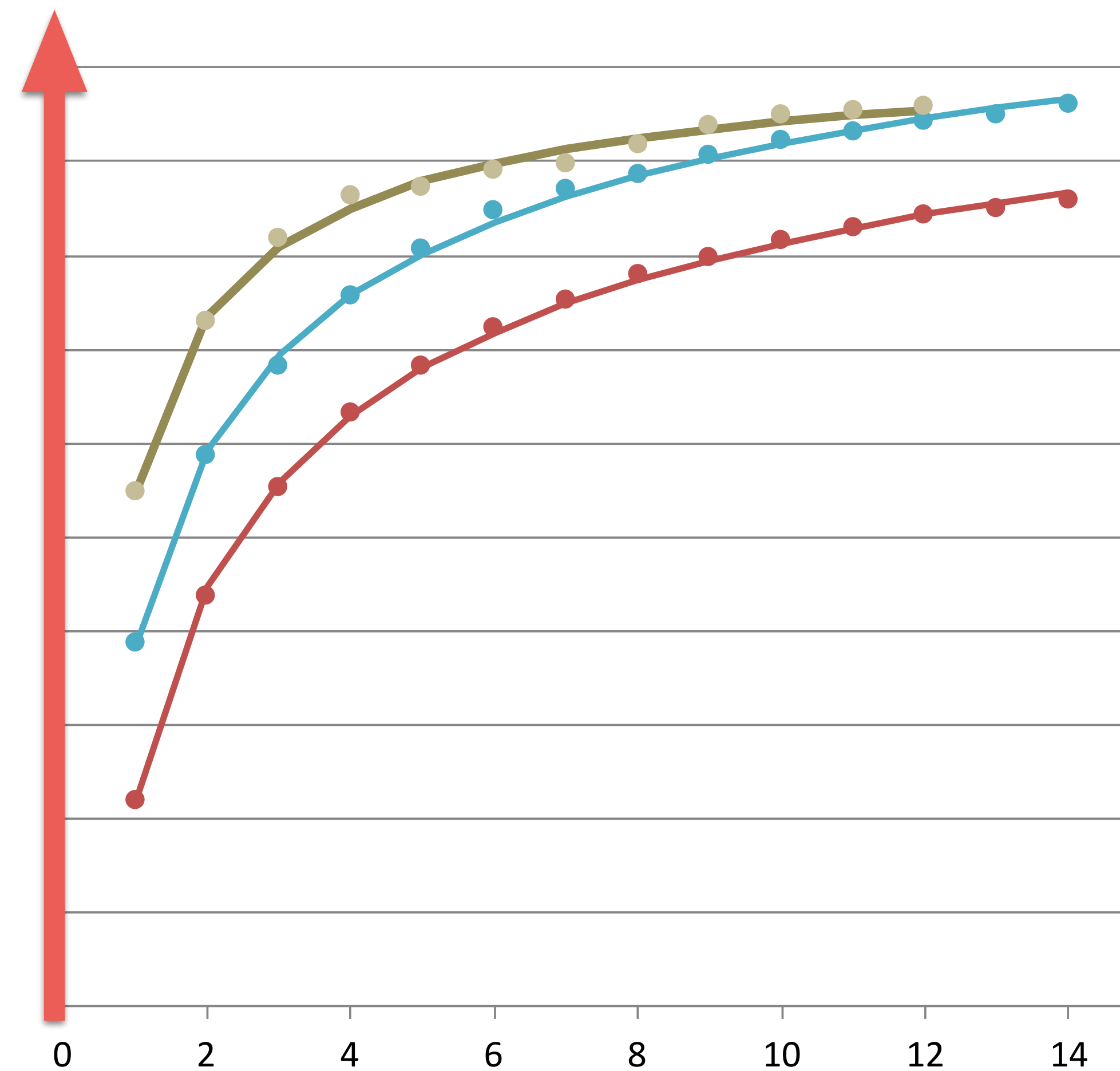
describe
30 sec } free-view
10 sec

BubbleView participants



More involved tasks lead to better clicks

Similarity between clicks and fixations



describe
30 sec
10 sec

↑ engagement

BubbleView participants



Take-away #3:

Task time and bubble size interact to affect clicks.

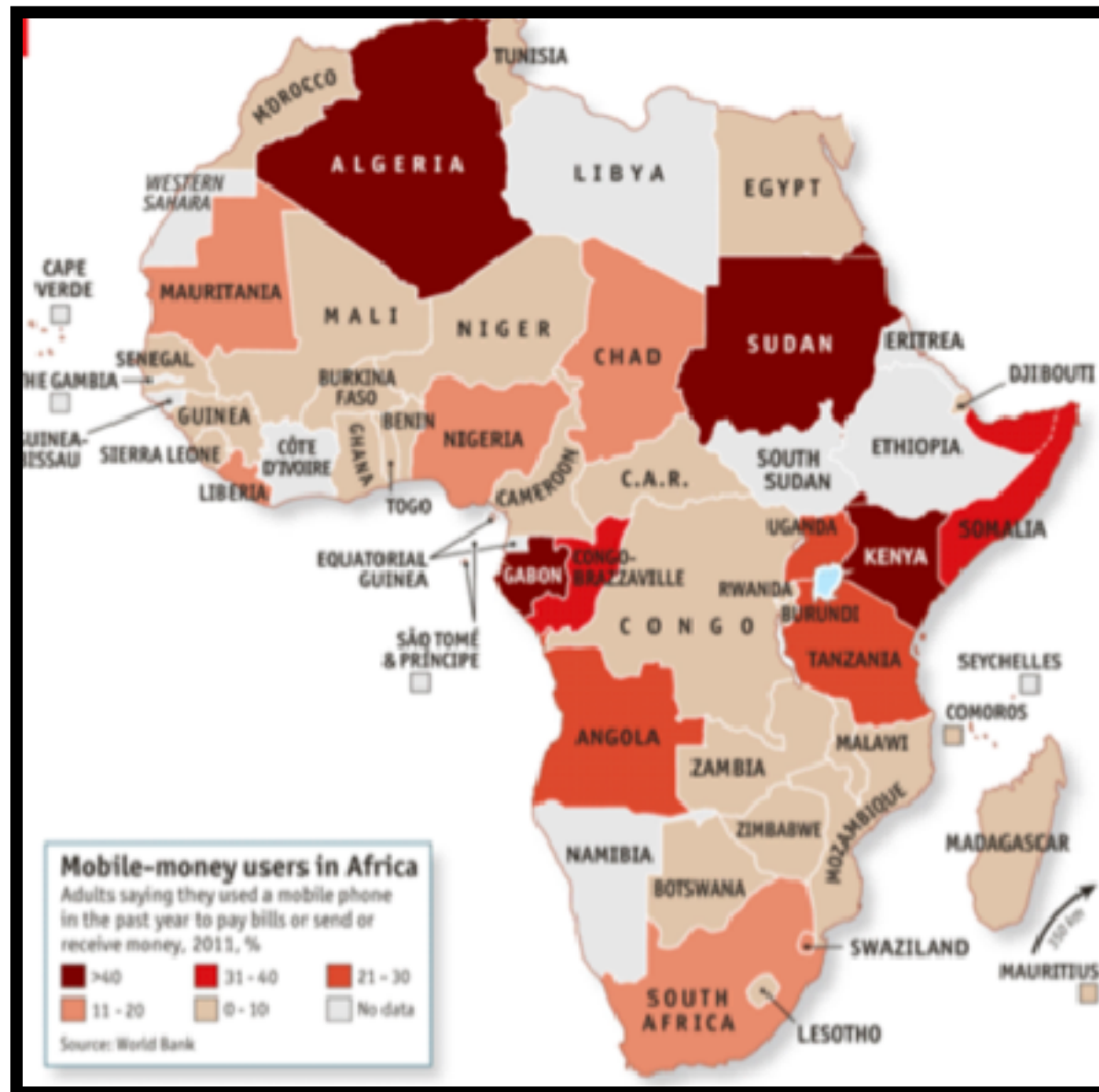


Performance is stable across bubble sizes

32 pix

24 pix

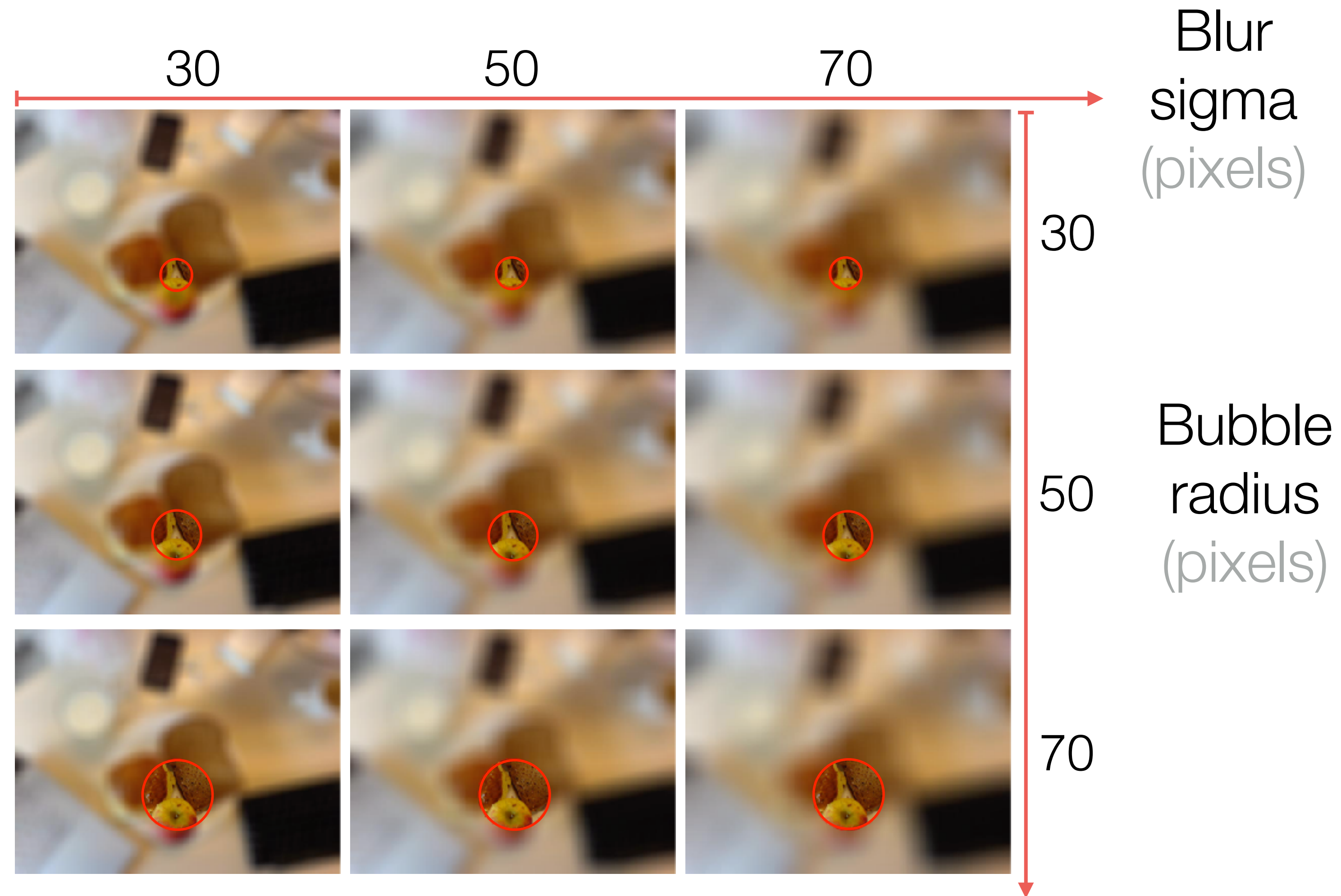
16 pix



x 1.5 clicks

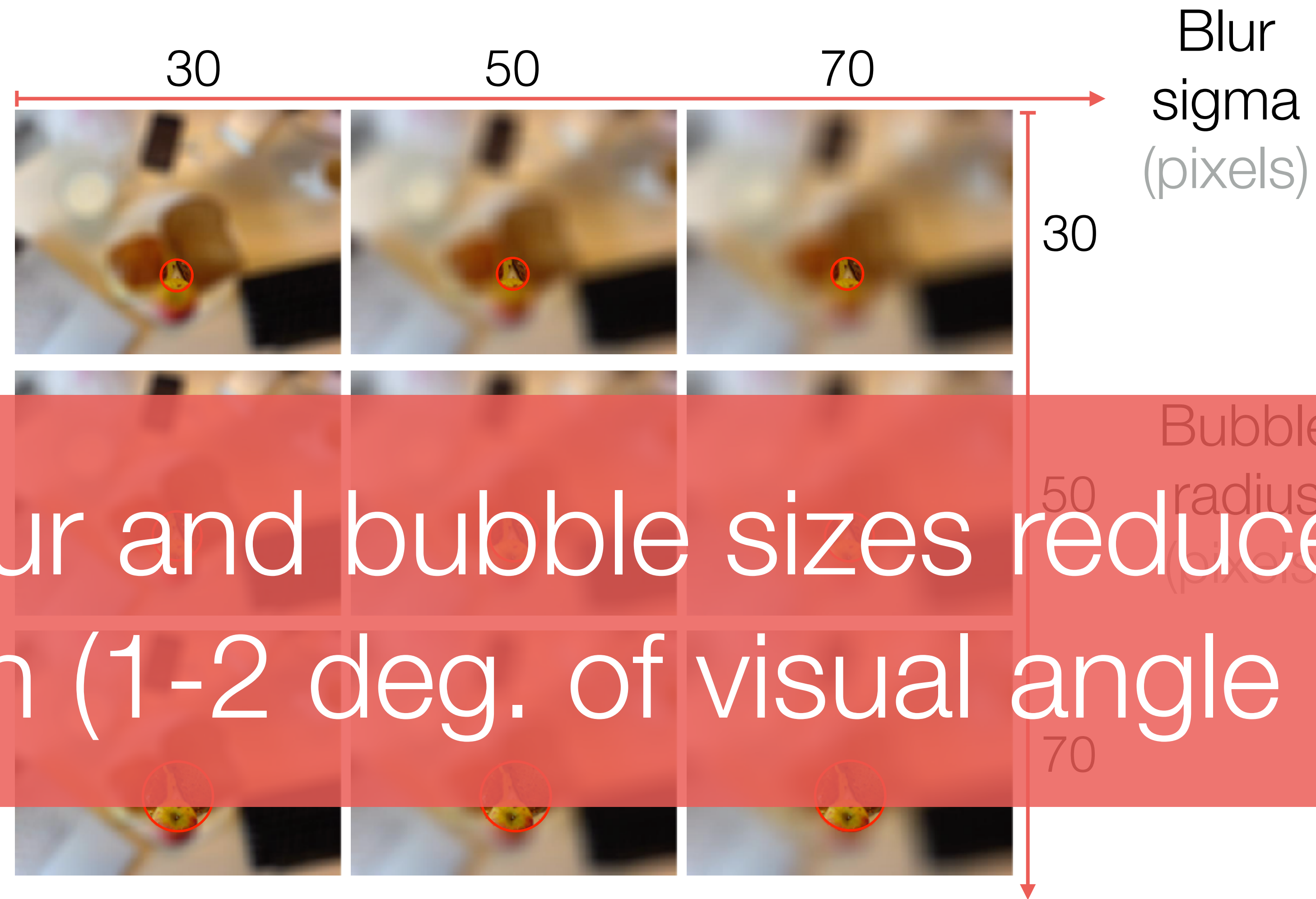


Blur affects clicks more than bubble size





Blur affects clicks more than bubble size

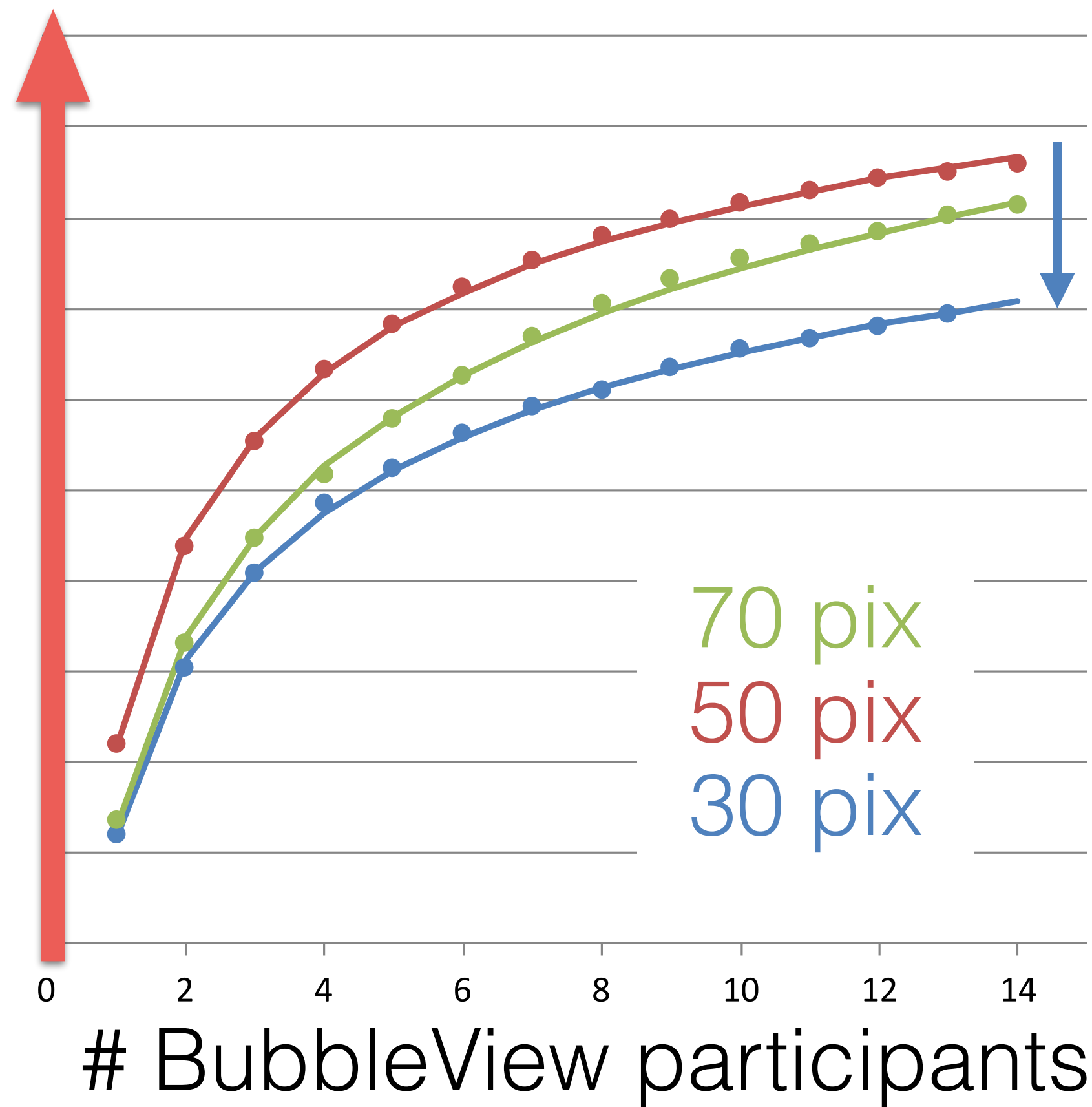


Task time and bubble size interact

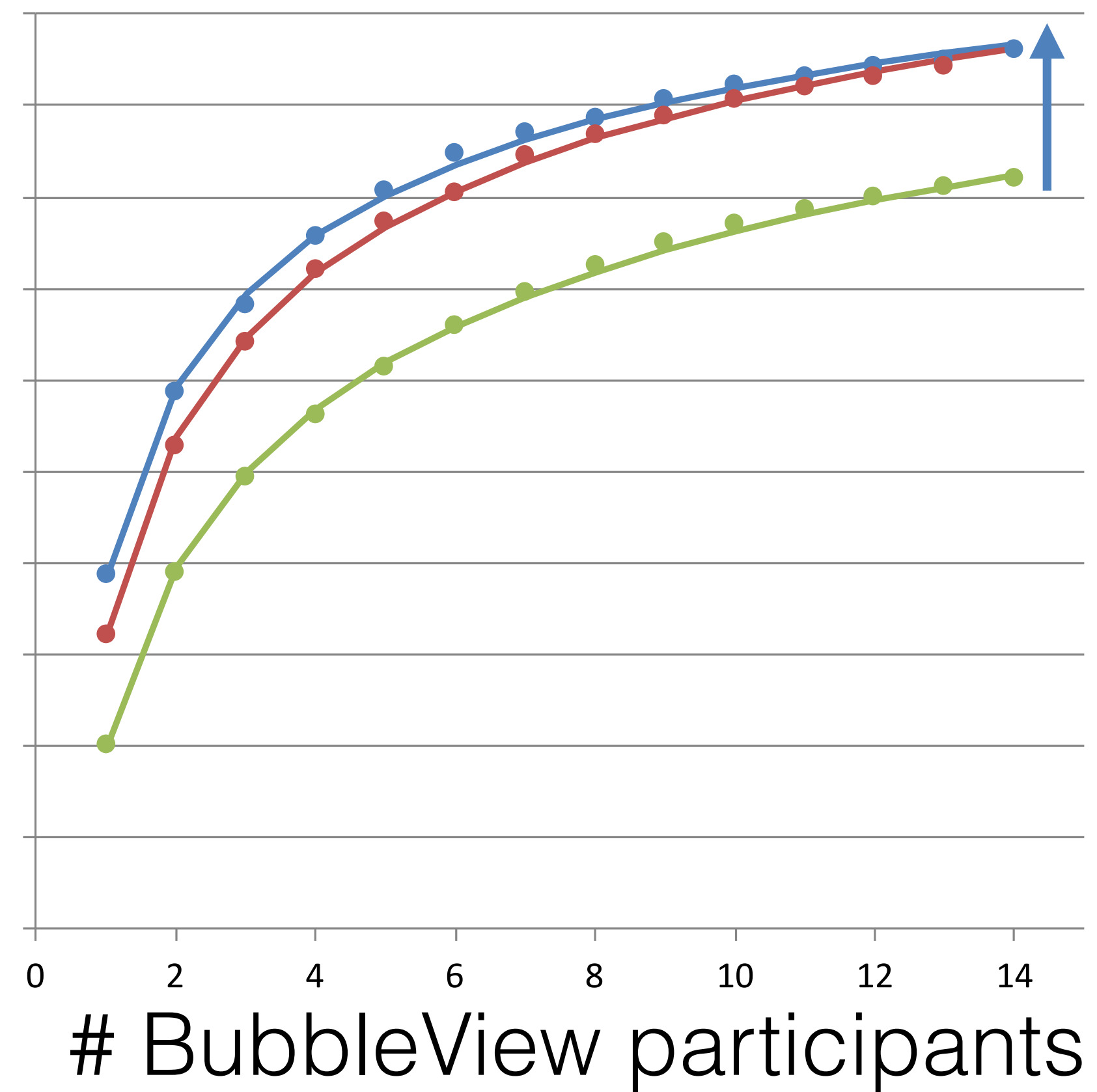
Similarity between clicks and fixations



10 sec task



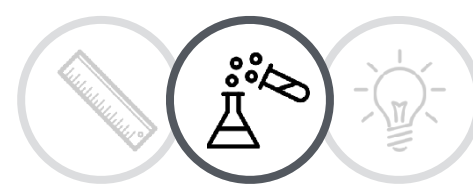
30 sec task



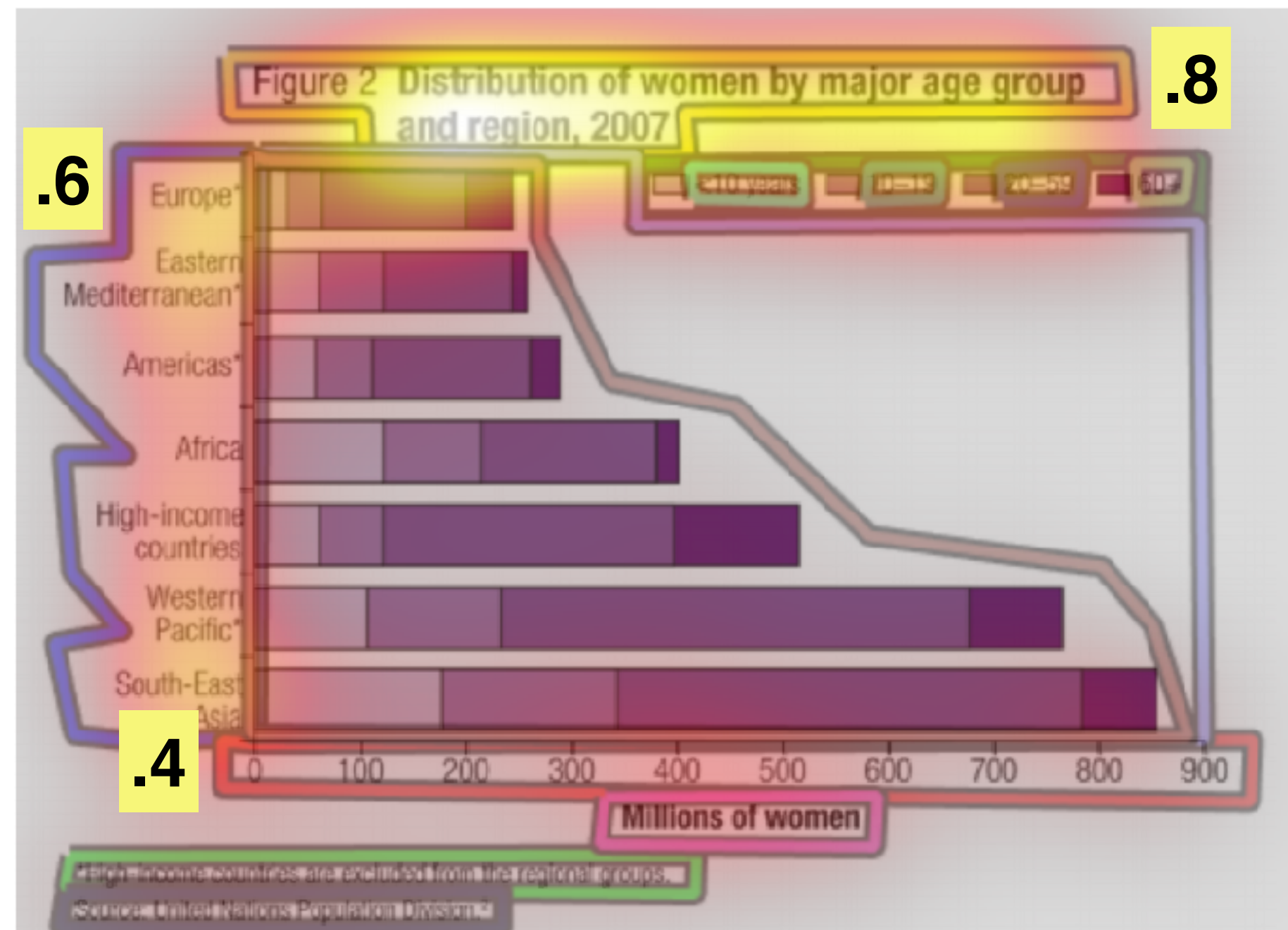


Take-away #4:

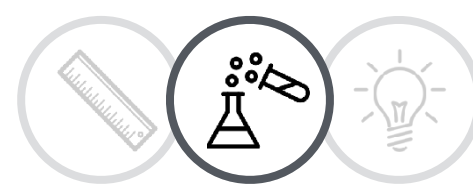
BubbleView can be used to rank image elements by importance.



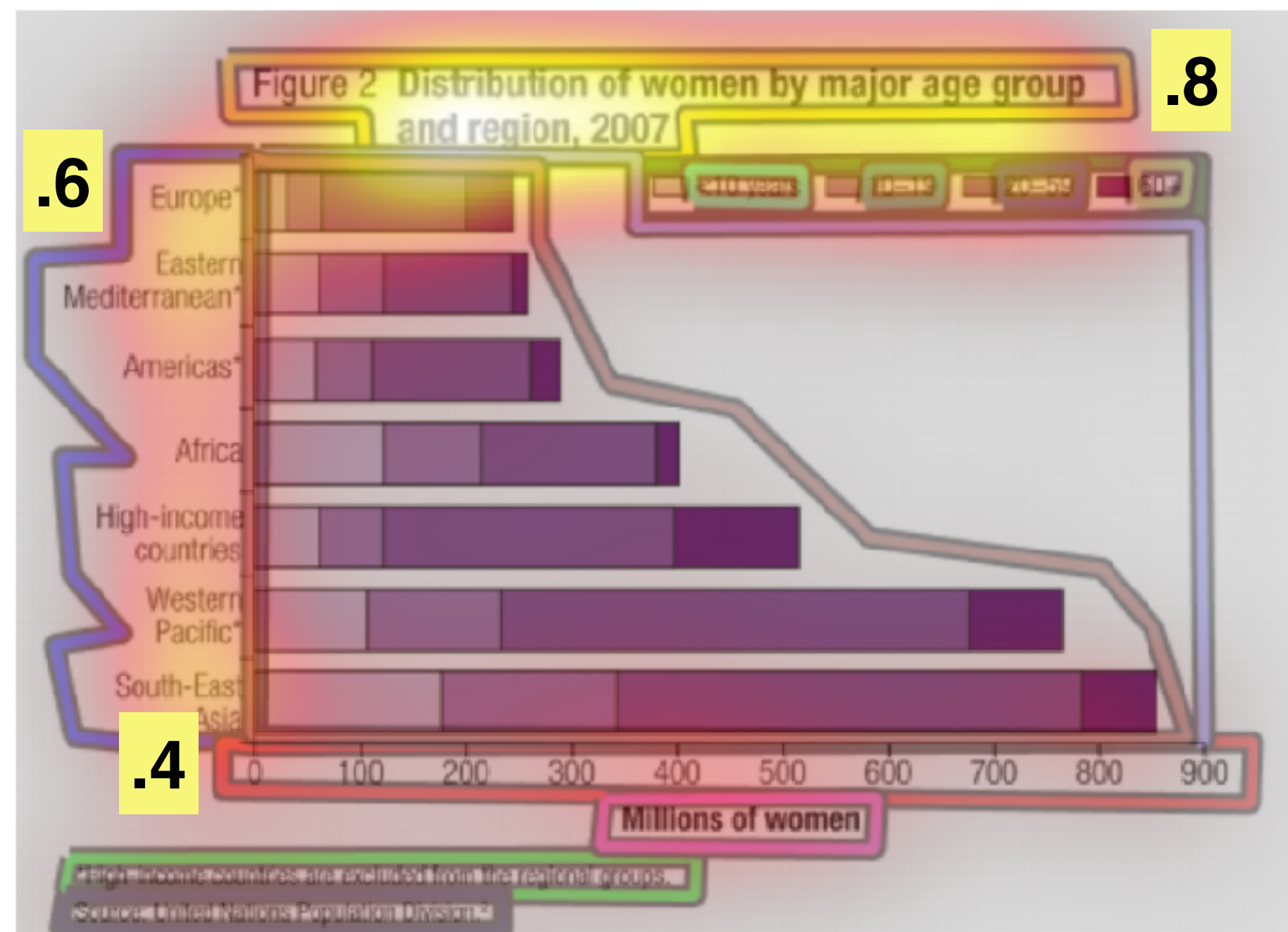
Ranking elements by importance



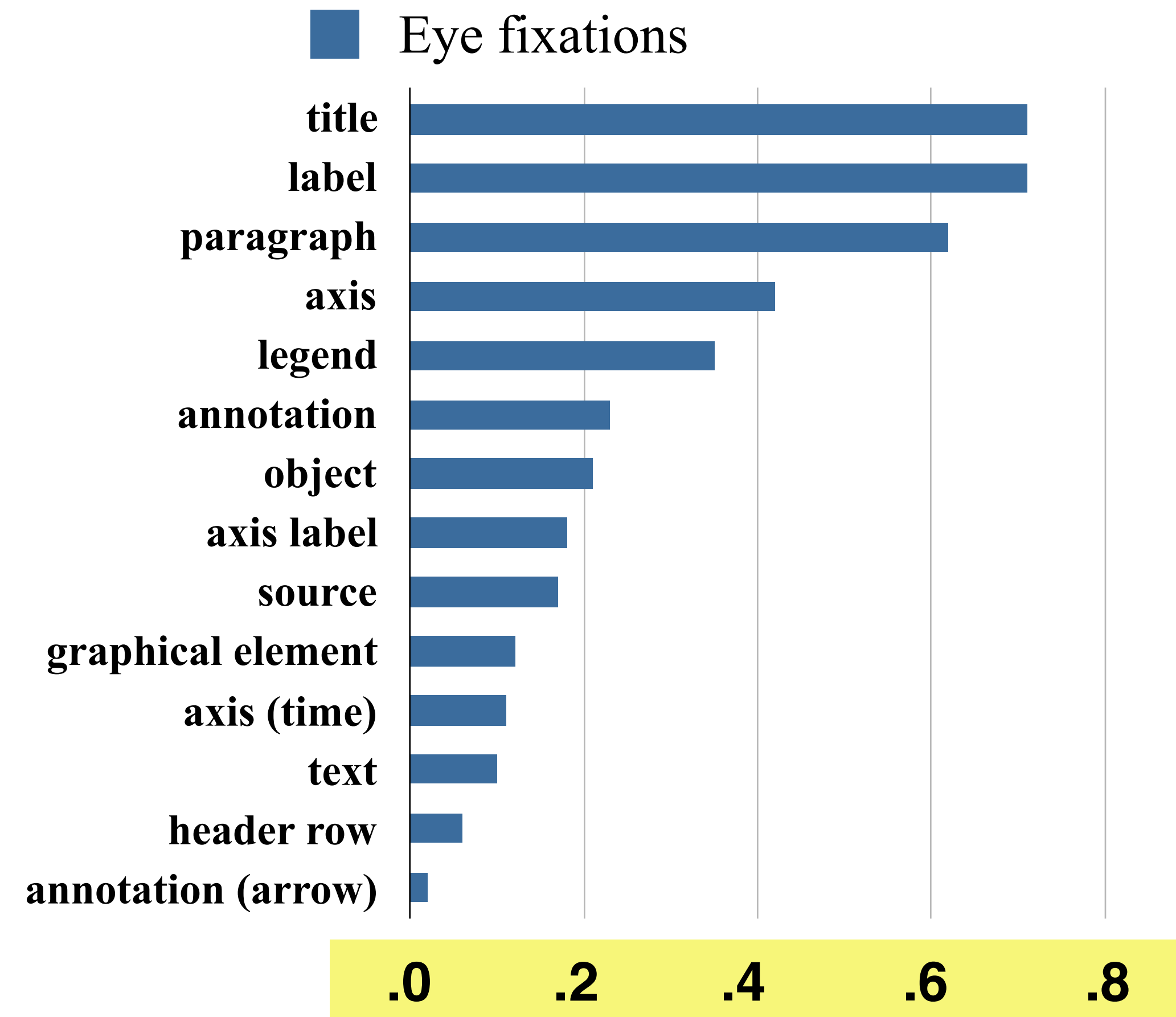
- (1) axis
- (2) axis
- (3) text (paragraph)
- (4) text (source)
- (5) text (axis label)
- (6) text (title)
- (7) legend
- (8) data
- (9) data (bars)
- (10) text (label)
- (11) text (label)
- (12) text (label)
- (13) text (label)



Ranking elements by importance

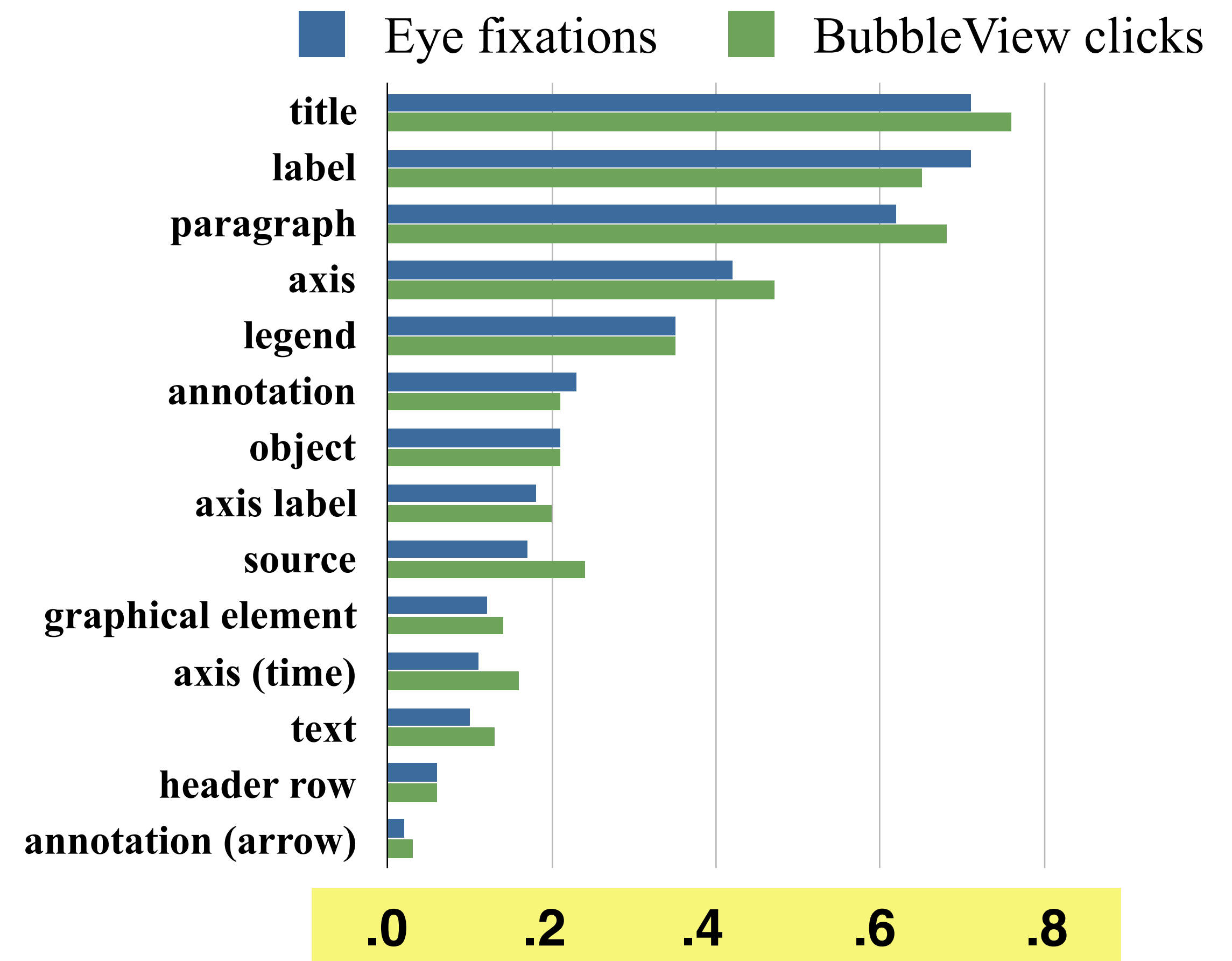
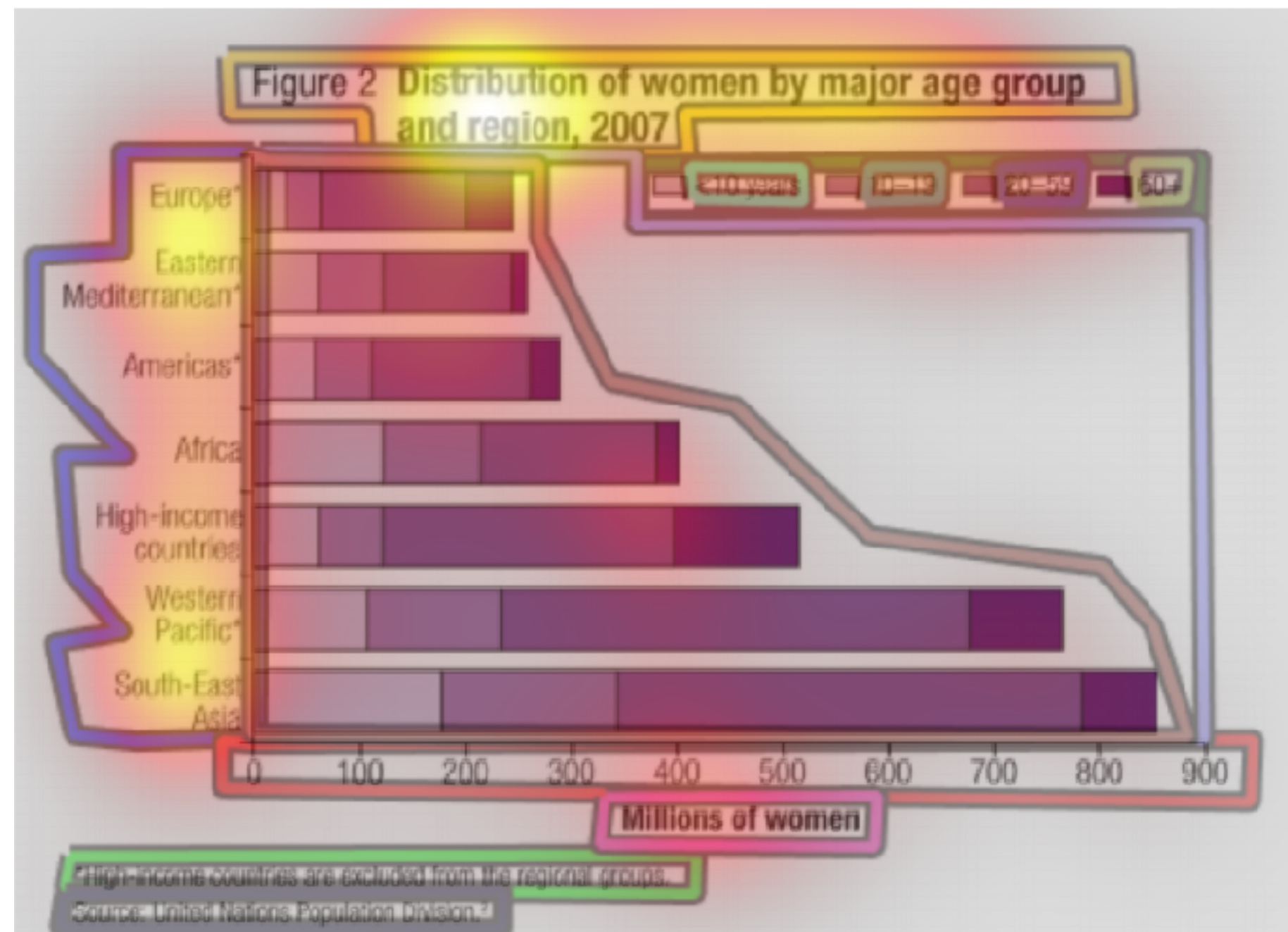


- (1) axis
- (2) axis
- (3) text (paragraph)
- (4) text (source)
- (5) text (axis label)
- (6) text (title)
- (7) legend
- (8) data
- (9) data (bars)
- (10) text (label)
- (11) text (label)
- (12) text (label)
- (13) text (label)



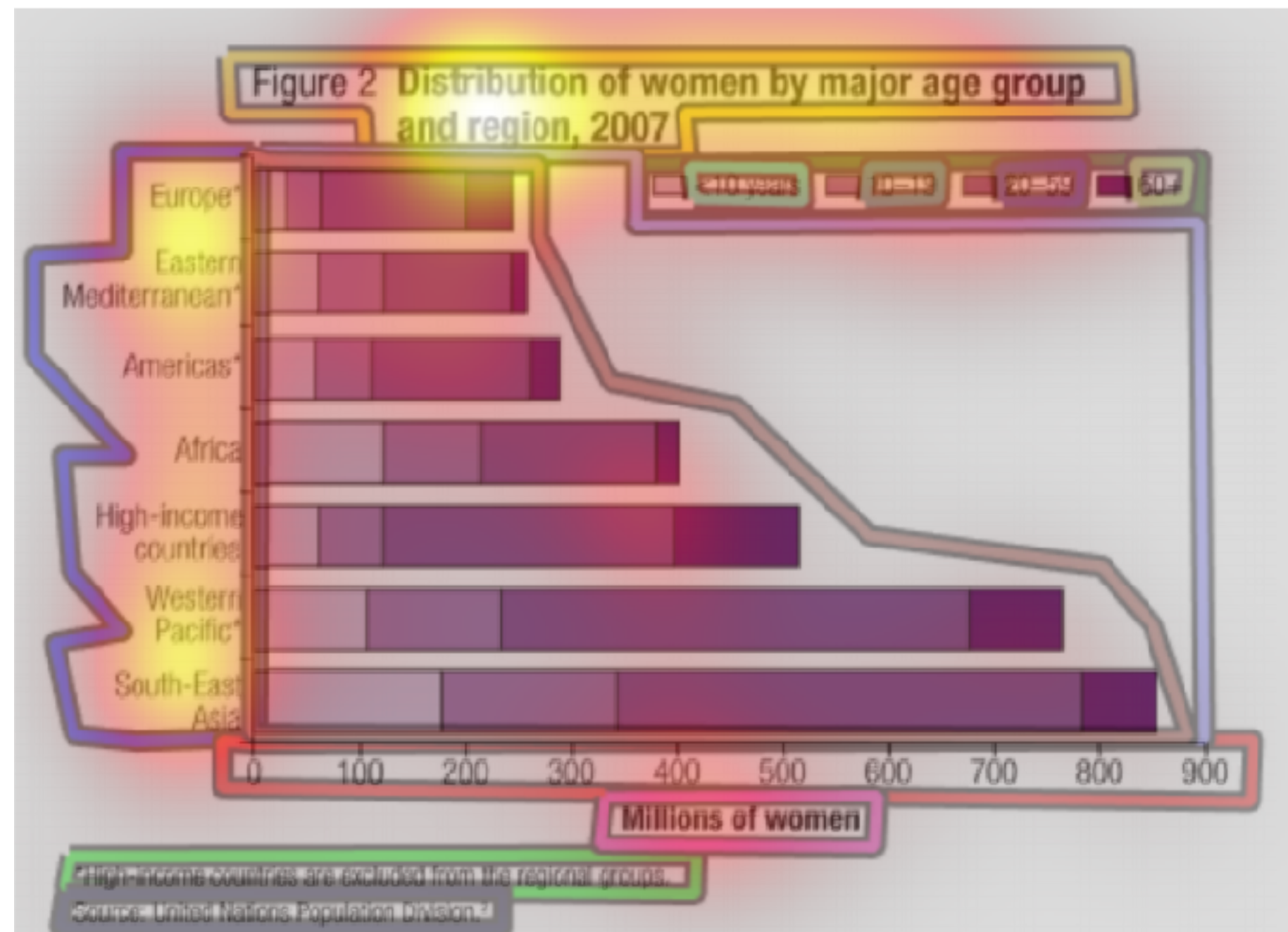


Ranking elements by importance

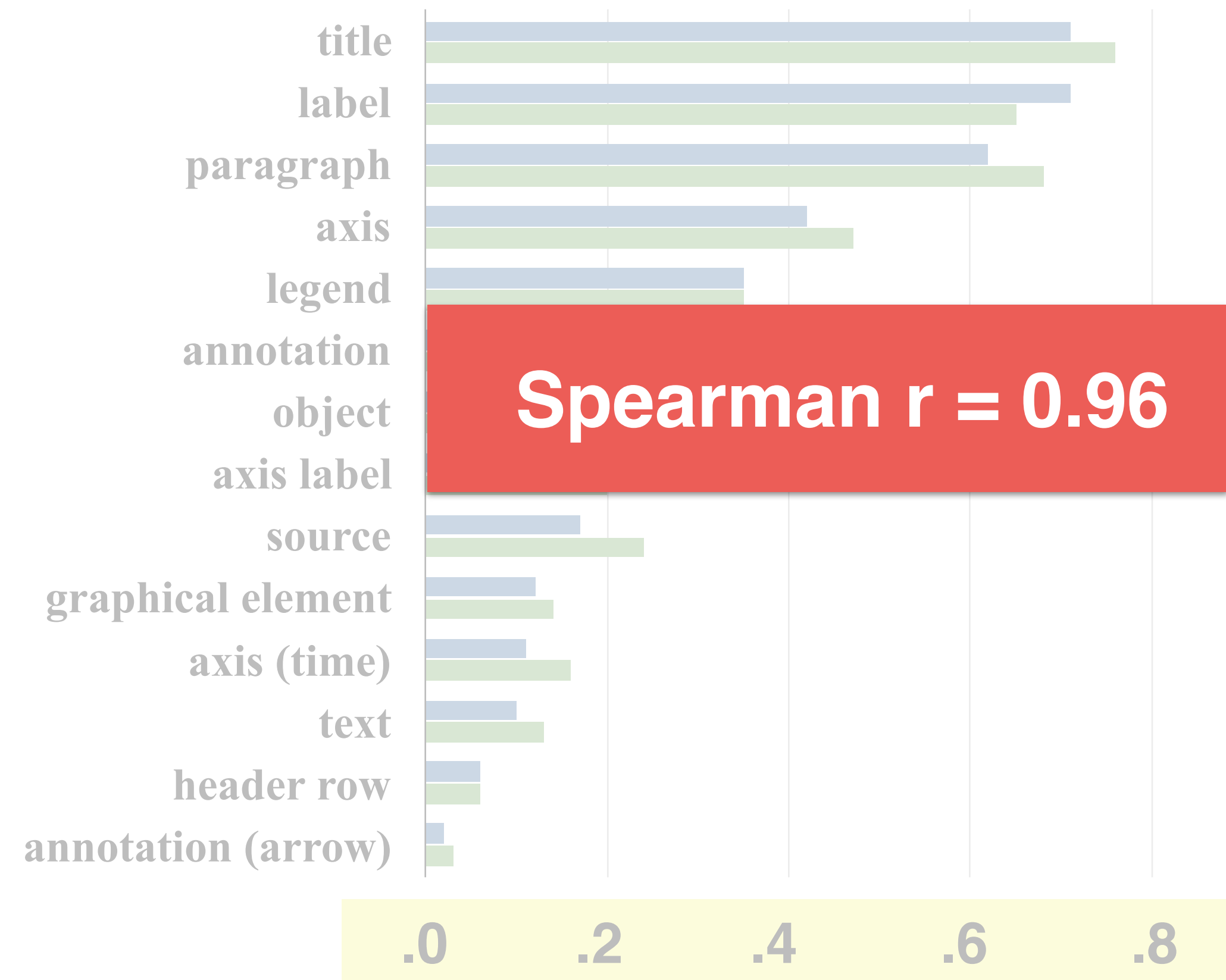




Ranking elements by importance

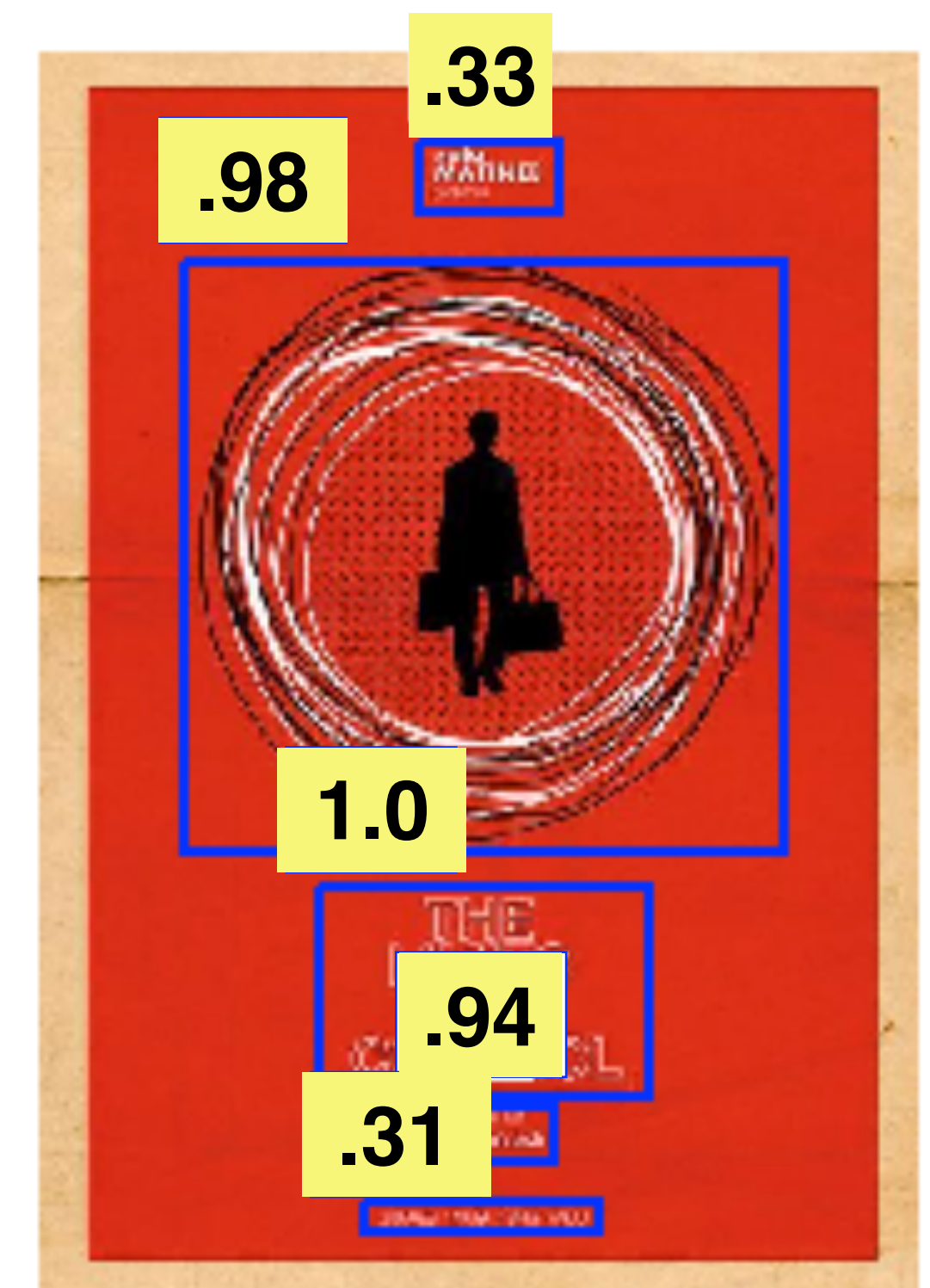
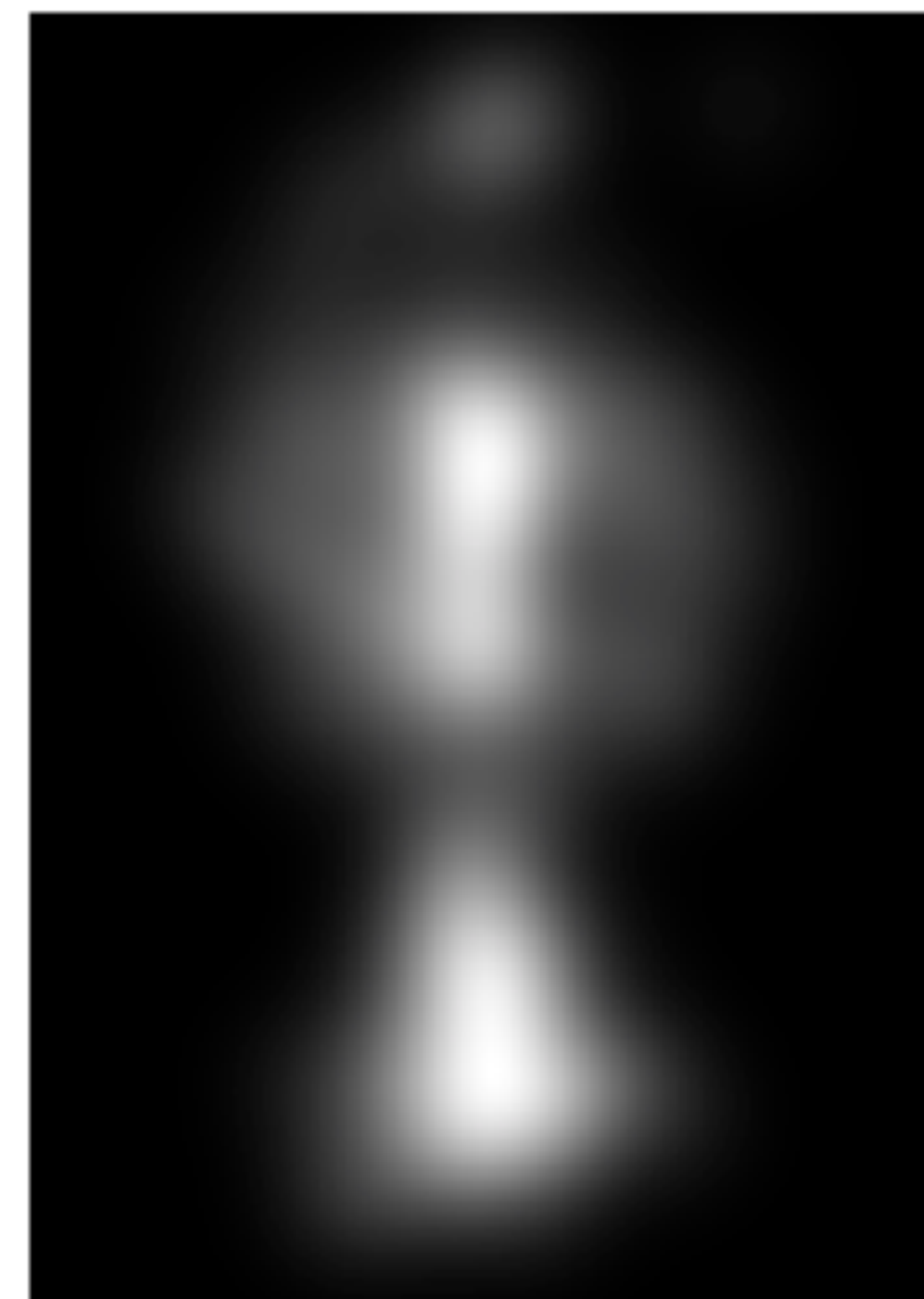
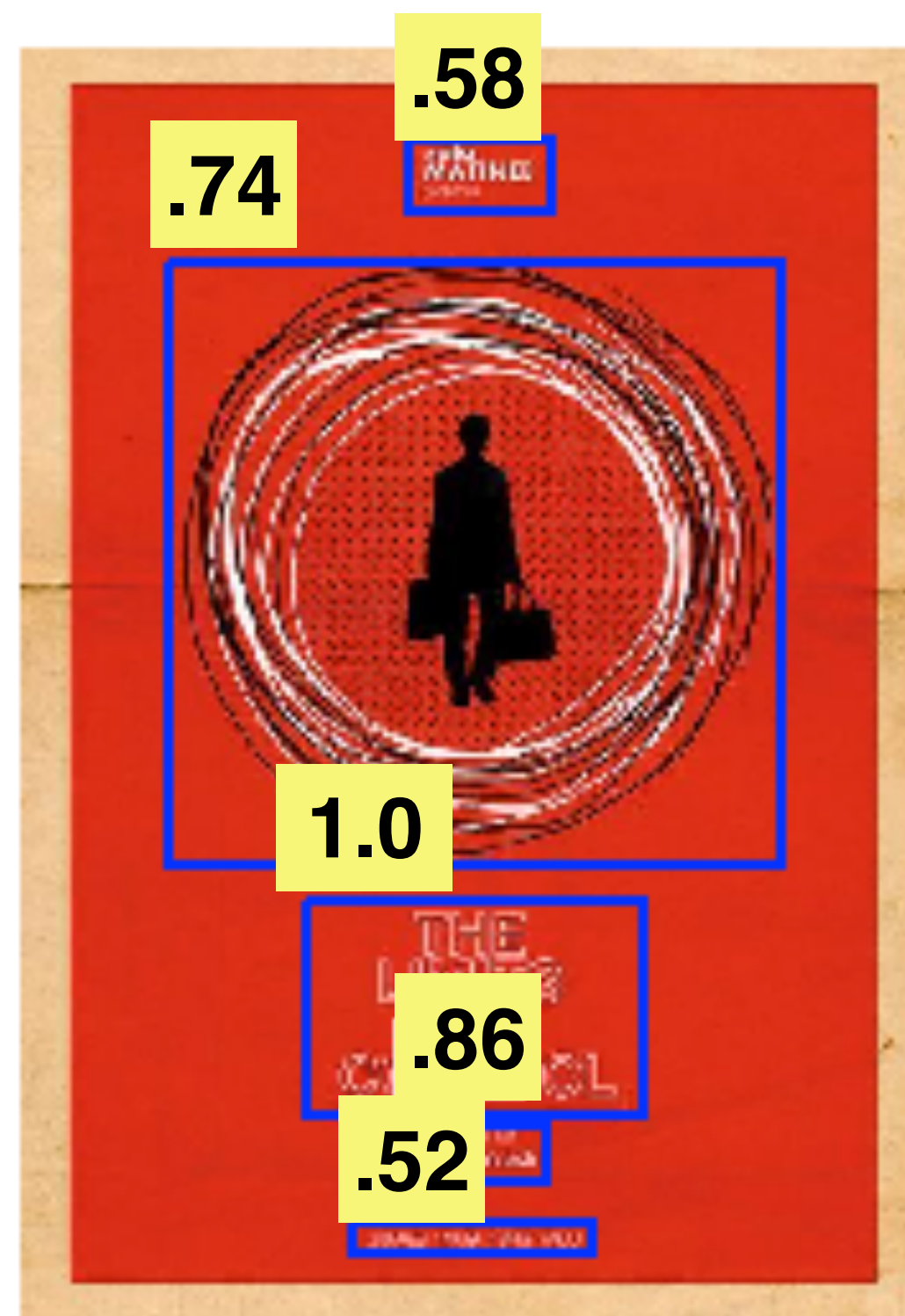
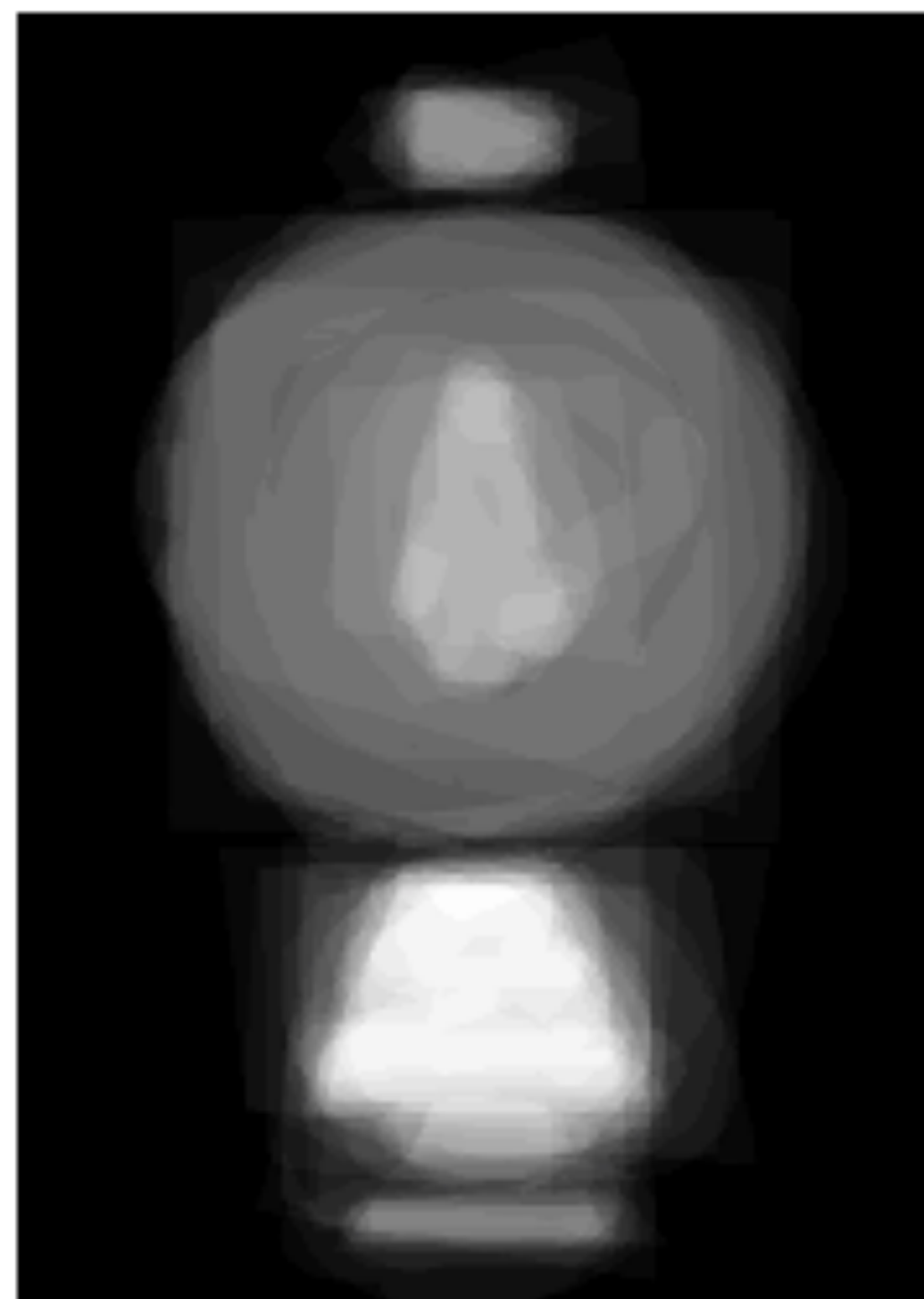


■ Eye fixations ■ BubbleView clicks





Ranking elements by importance

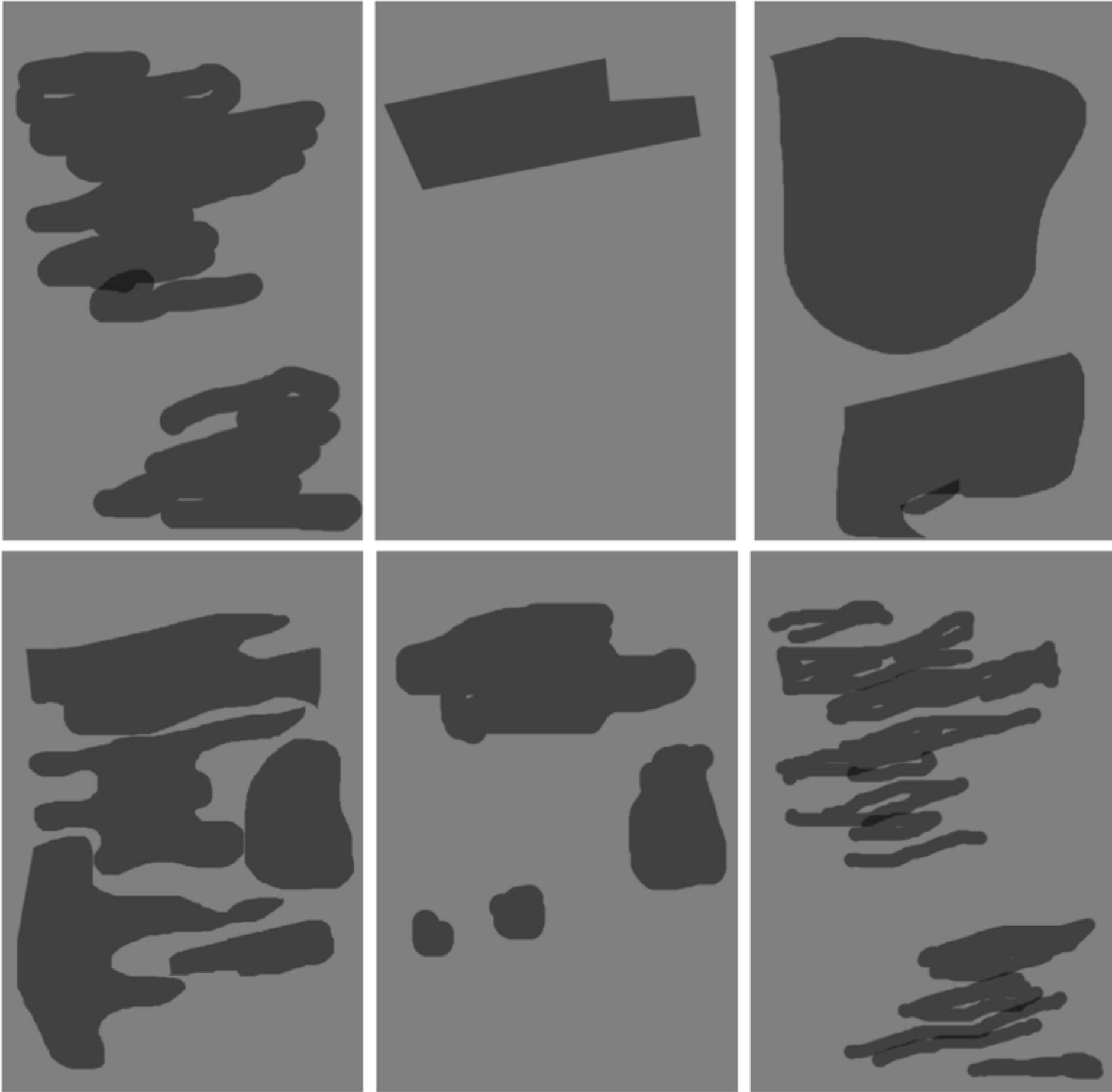


Spearman $r = 0.60$

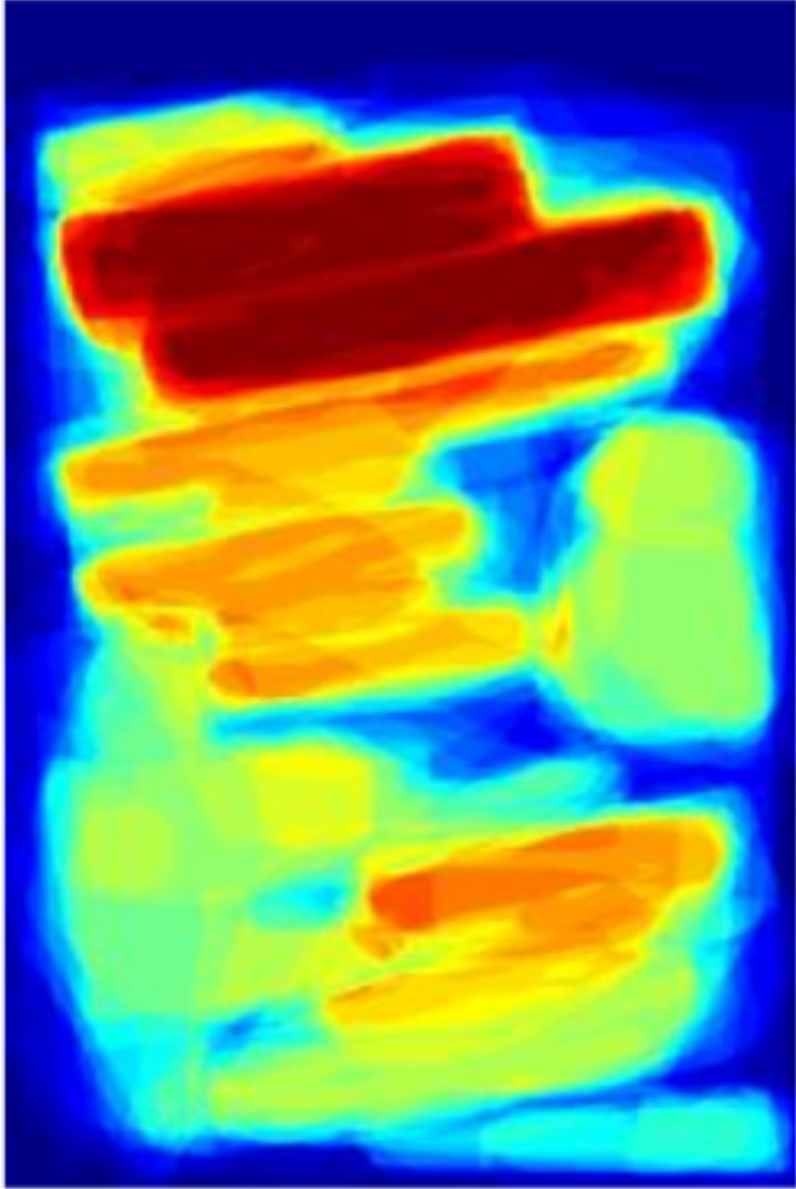
Another measurement of importance

Crowd annotations

Graphic design



Avg. annotation





Design choice: collecting importance

Fixations



“unconscious”

explorative

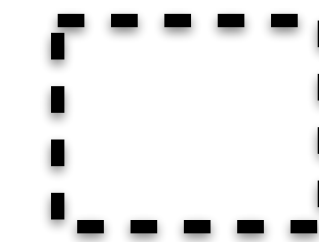
Clicks



conscious

explorative

Annotations



conscious

constrained



BubbleView measures importance

Fixations



with free-viewing

Clicks



with free-viewing

Clicks



with description

“Saliency”

Intentionality

“Importance”

Effort, task time, consistency



Evaluation Tools



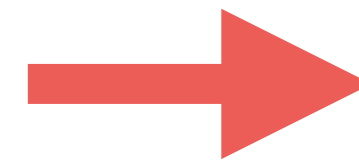
Experimental Results



Future Applications

Retargeting & Thumbnailing

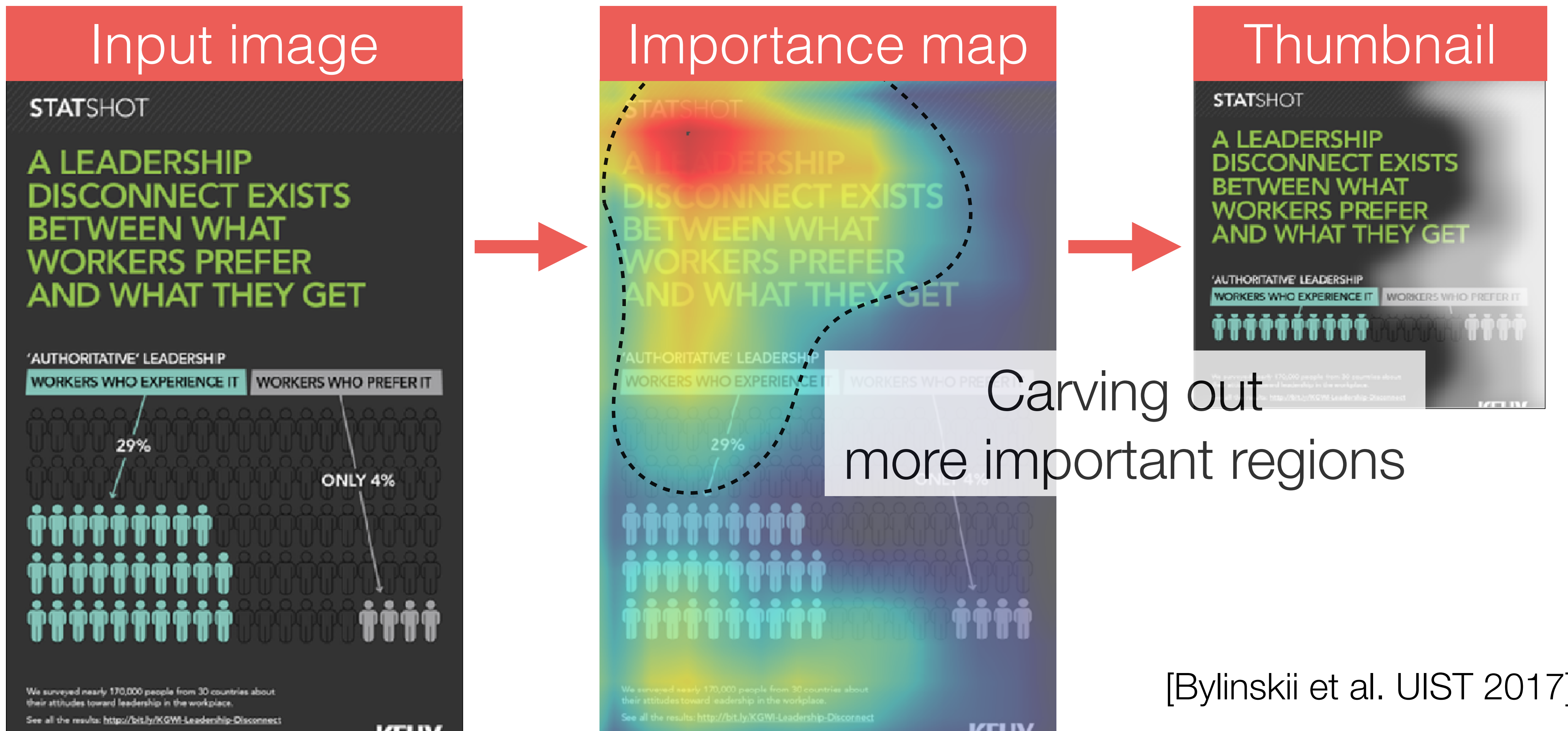
Input image



Importance map

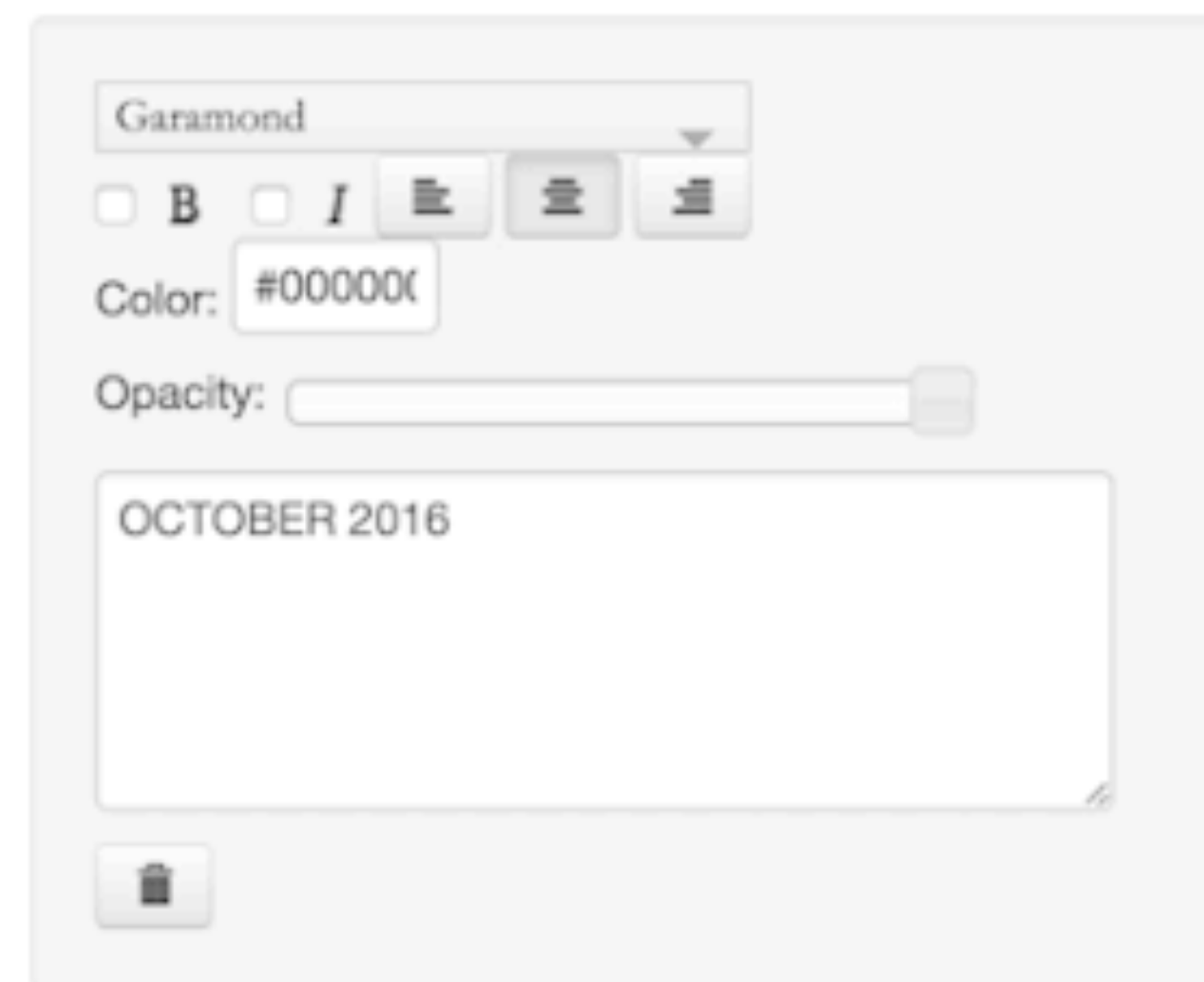


Retargeting & Thumbnailing





Prediction of Visual Importance



Providing real time feedback
based on importance predictions

[\[visimportance.csail.mit.edu\]](http://visimportance.csail.mit.edu)



Task-specific attentional data

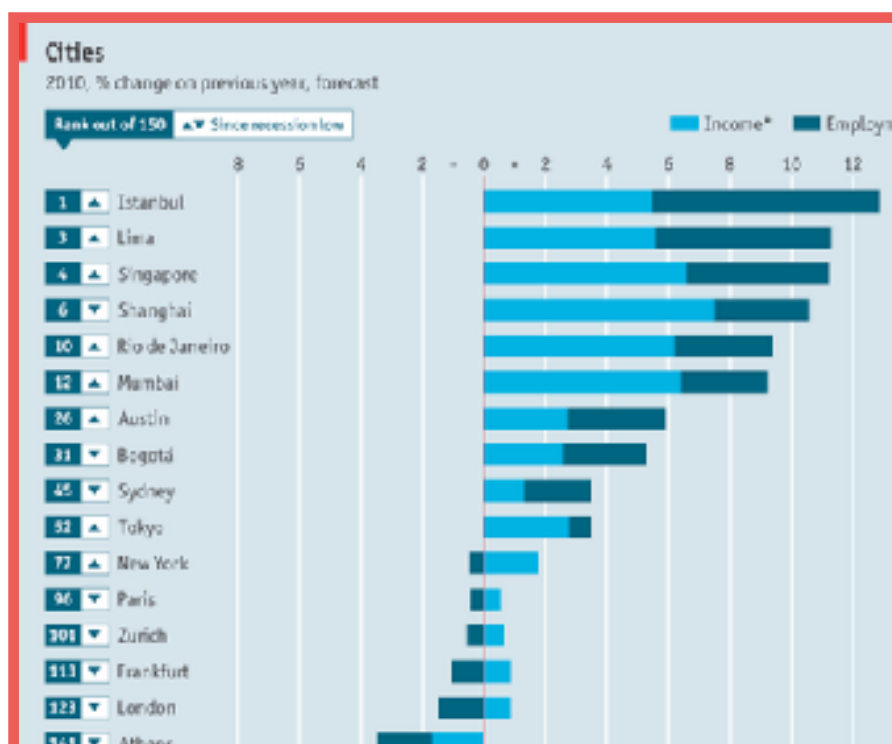
Which city is ranked first (find an extremum)?



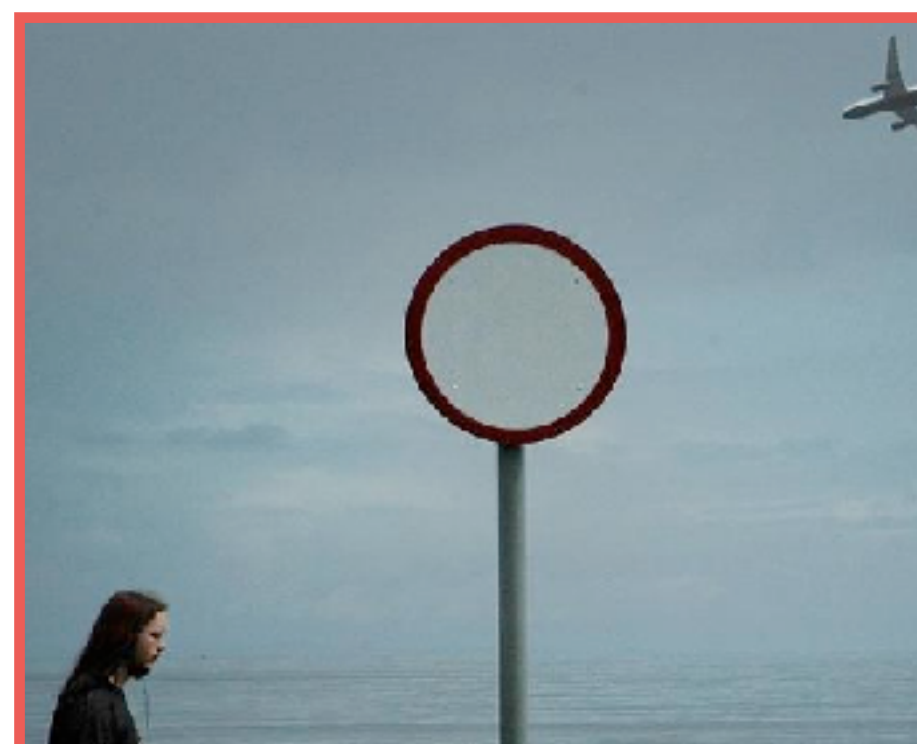
Visual Question Answering



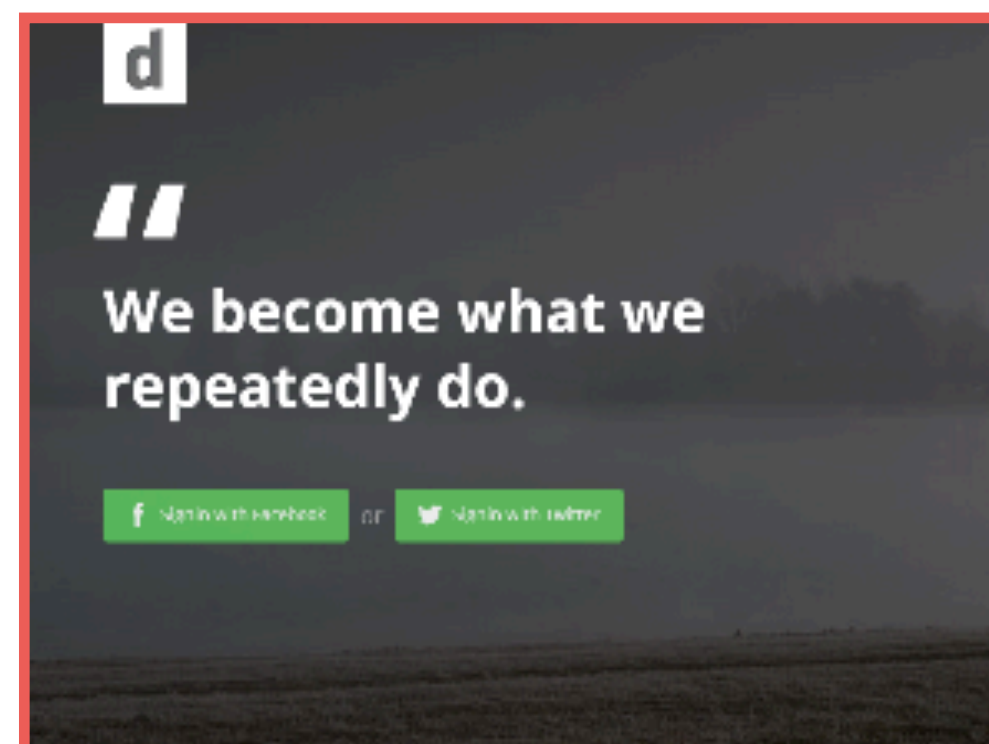
Other image domains & tasks



Visualization



Natural scene



Webpage



Graphic design

Tasks

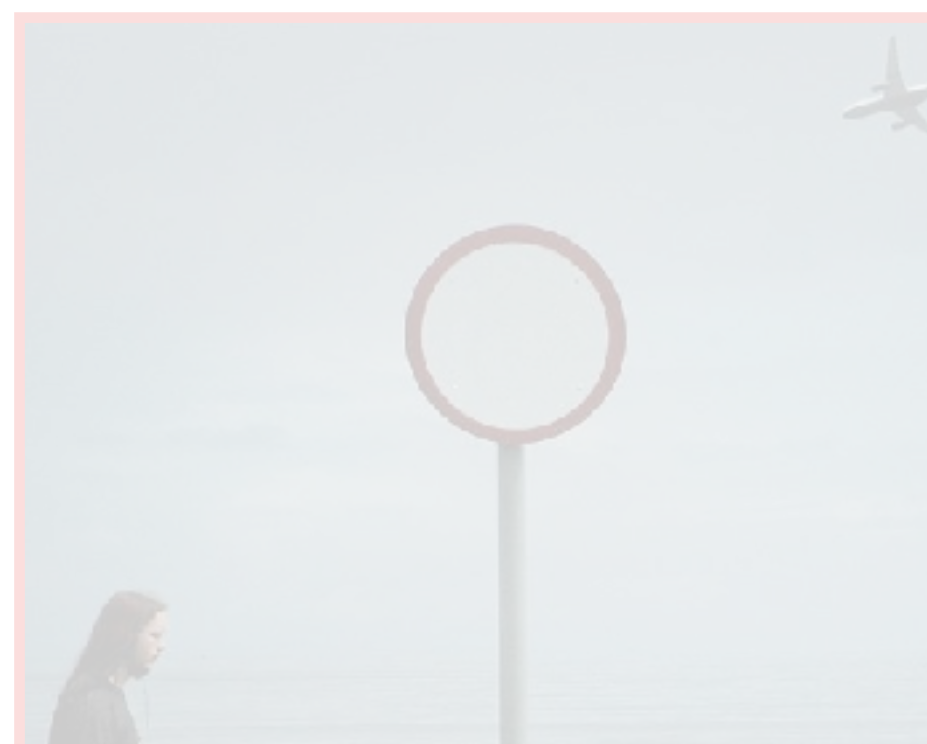
Free-viewing
Description



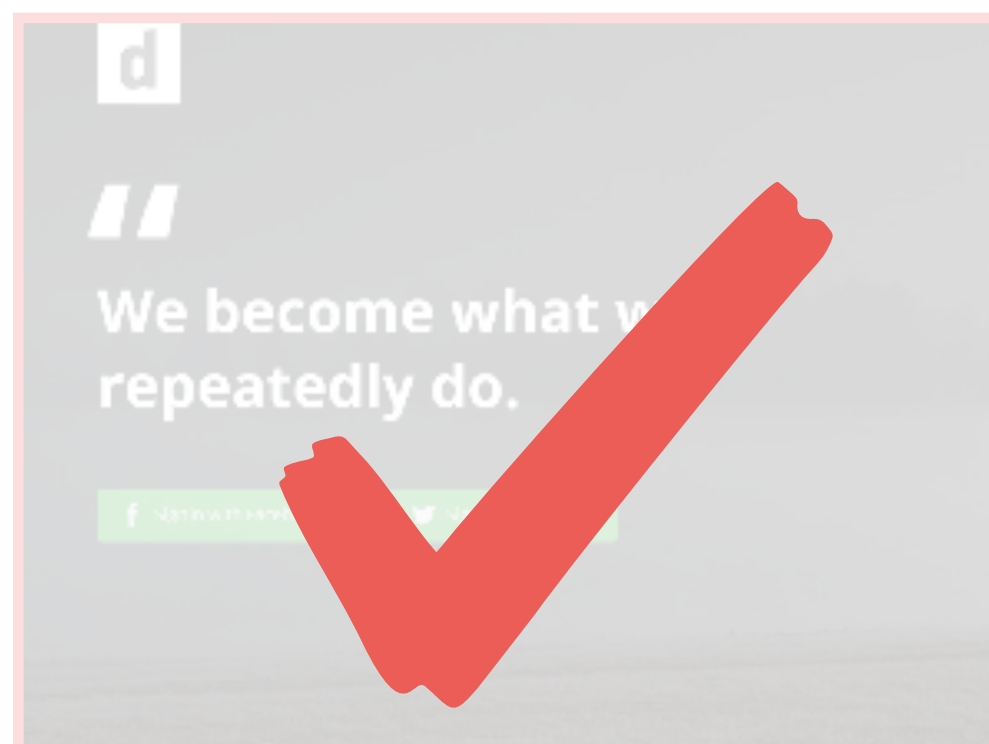
Other image domains & tasks



Visualization



Natural scene



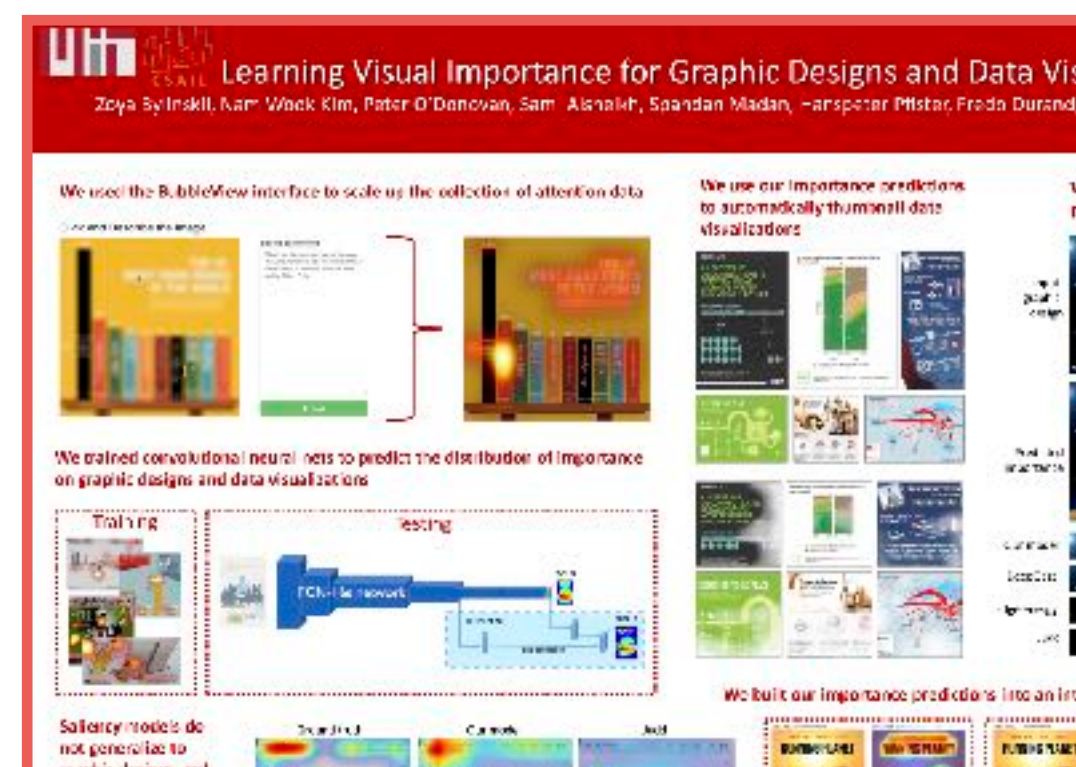
Webpage



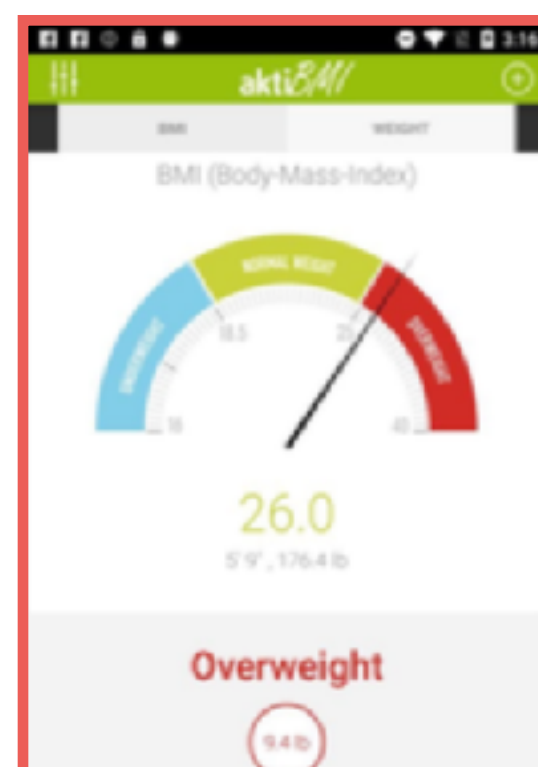
Graphic design

Tasks

Free-viewing
Description



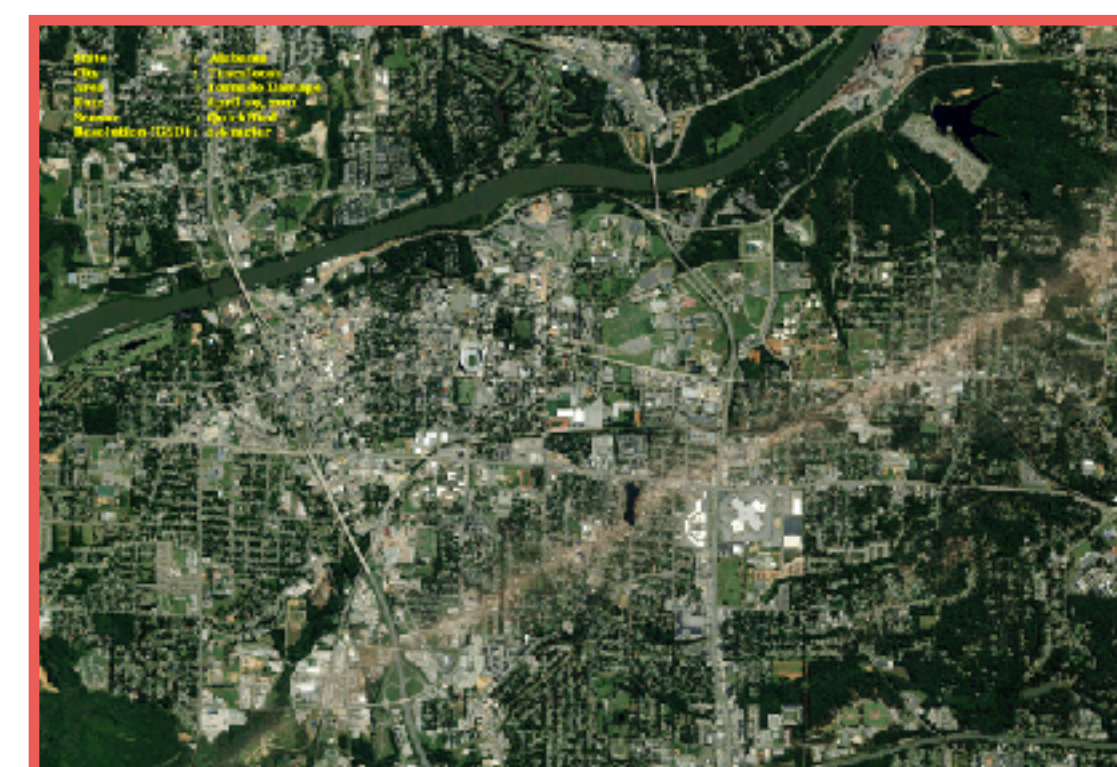
Posters



Mobile



Medical images



Satellite images

Visual search
Q & A
Analysis tasks

0 clicks 0 characters

Describe the image in as much detail as possible...

<http://bubbleview.namwkim.org>



Paper



Code



Demo

>> 2x

NEXT

