Introducing the TurkEyes Toolbox
UIs for crowdsourcing attention without an eye tracker

Anelise Newman
Barry McNamara
Camilo Fosco
Yun Bin Zhang

Pat Sukhum
Matthew Tancik
Nam Wook Kim
Zoya Bylinskii

The TurkEyes Tools

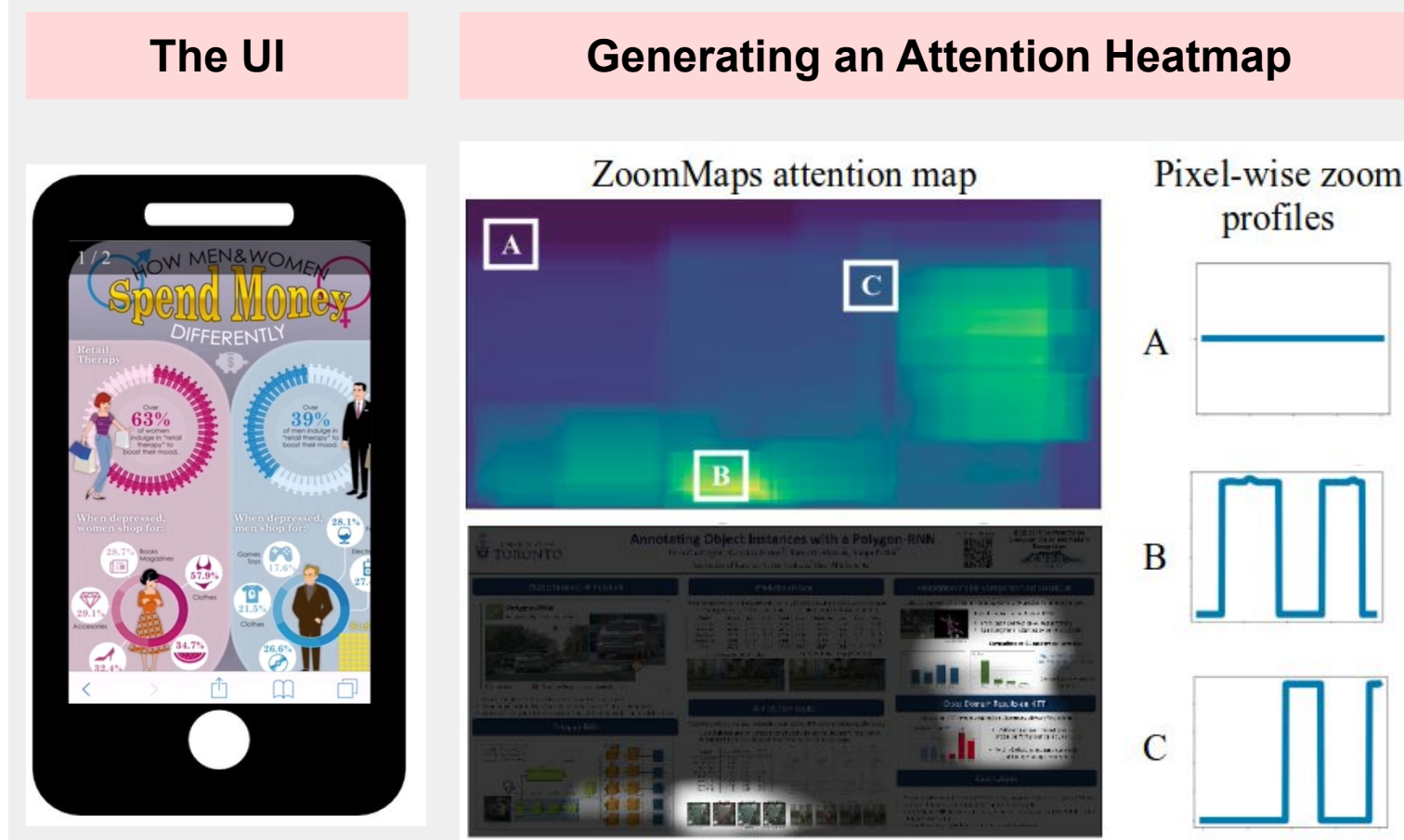We present a toolbox of **four user interfaces** for **crowdsourcing attention data**.

These interfaces **do not use eye tracking**; they instead rely on **interactions** with a computer or mobile phone that correlate with visual attention.
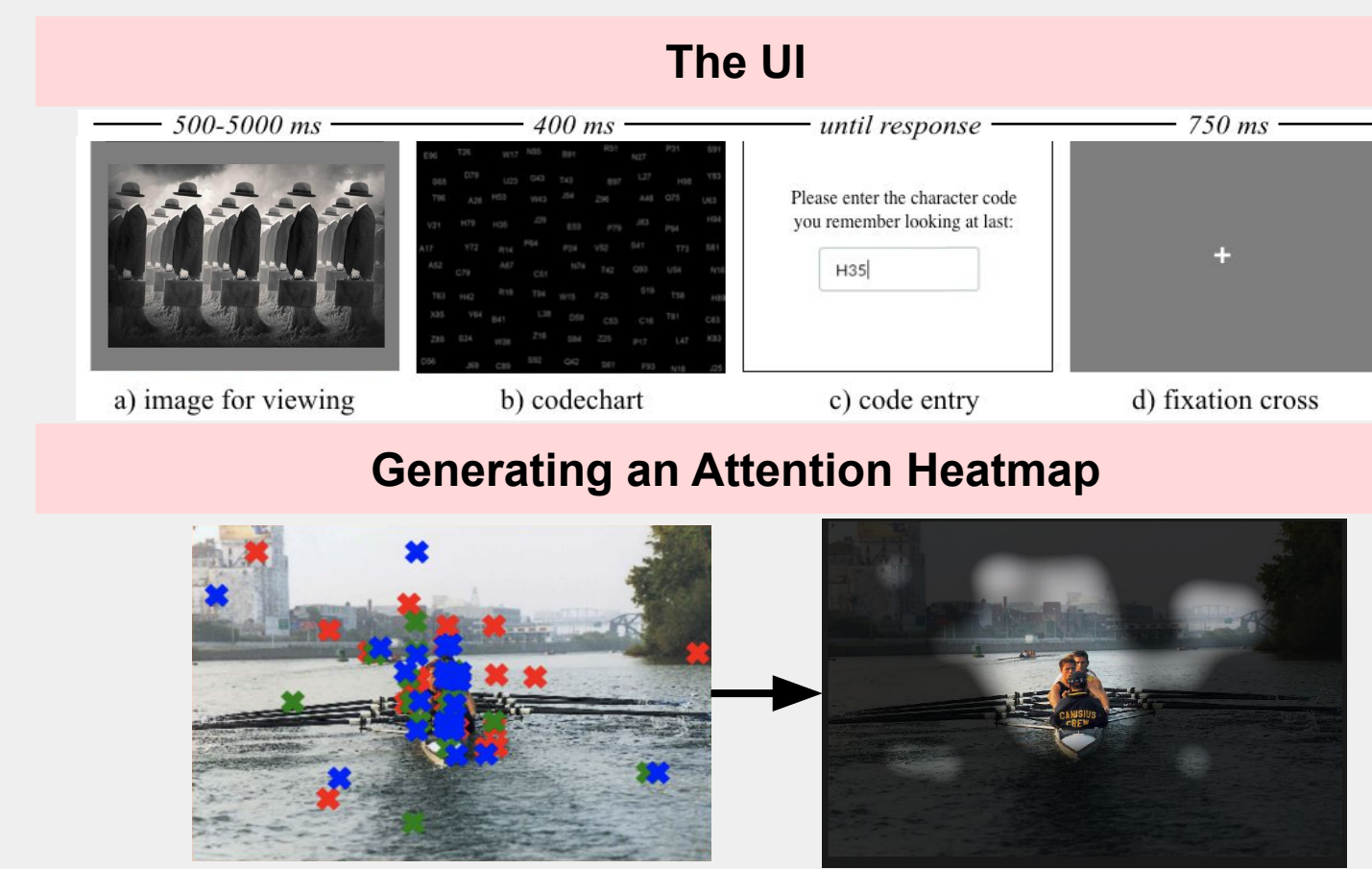
ZoomMaps (zoom-based)

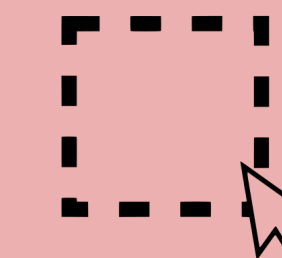*Participants use the pinch-zoom gesture on a phone to explore an image at multiple scales.*

The UI

Generating an Attention Heatmap

ZoomMaps attention map

Pixel-wise zoom profiles

A
B
C

CodeCharts (self-report)

Participants self-report where they gazed using a grid of three-letter codes.

The UI

500-5000 ms	400 ms	until response	750 ms
a) image for viewing	b) codechart	Please enter the character code you remember looking at last:	d) fixation cross
		c) code entry	

Generating an Attention Heatmap

ImportAnnots (annotation)

Participants paint over regions of a design that they consider important.

The UI

Generating a Heatmap
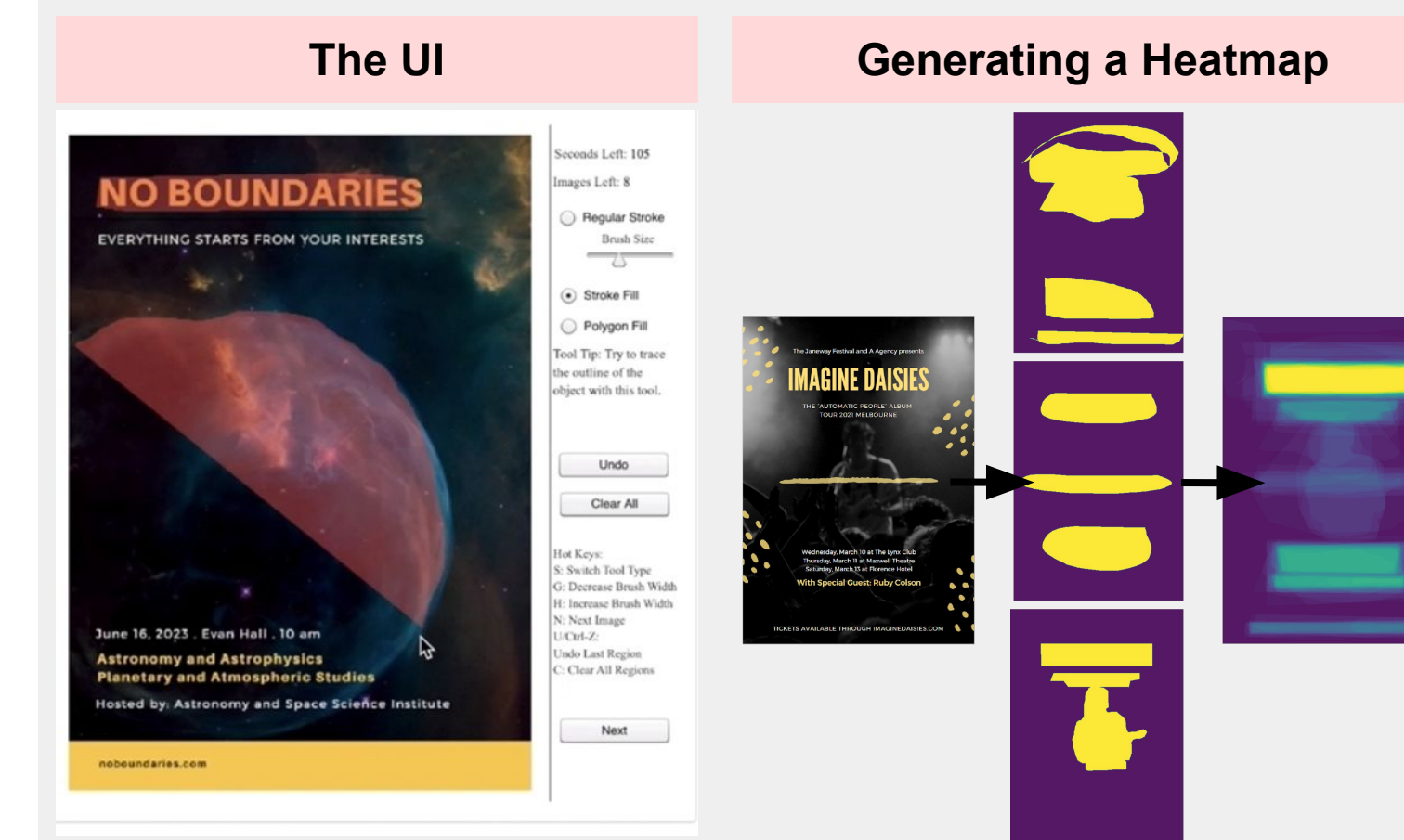
BubbleView (cursor-based)
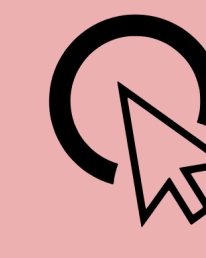
Participants click to unblur small bubble regions of a blurred image.

The UI

Generating a Heatmap
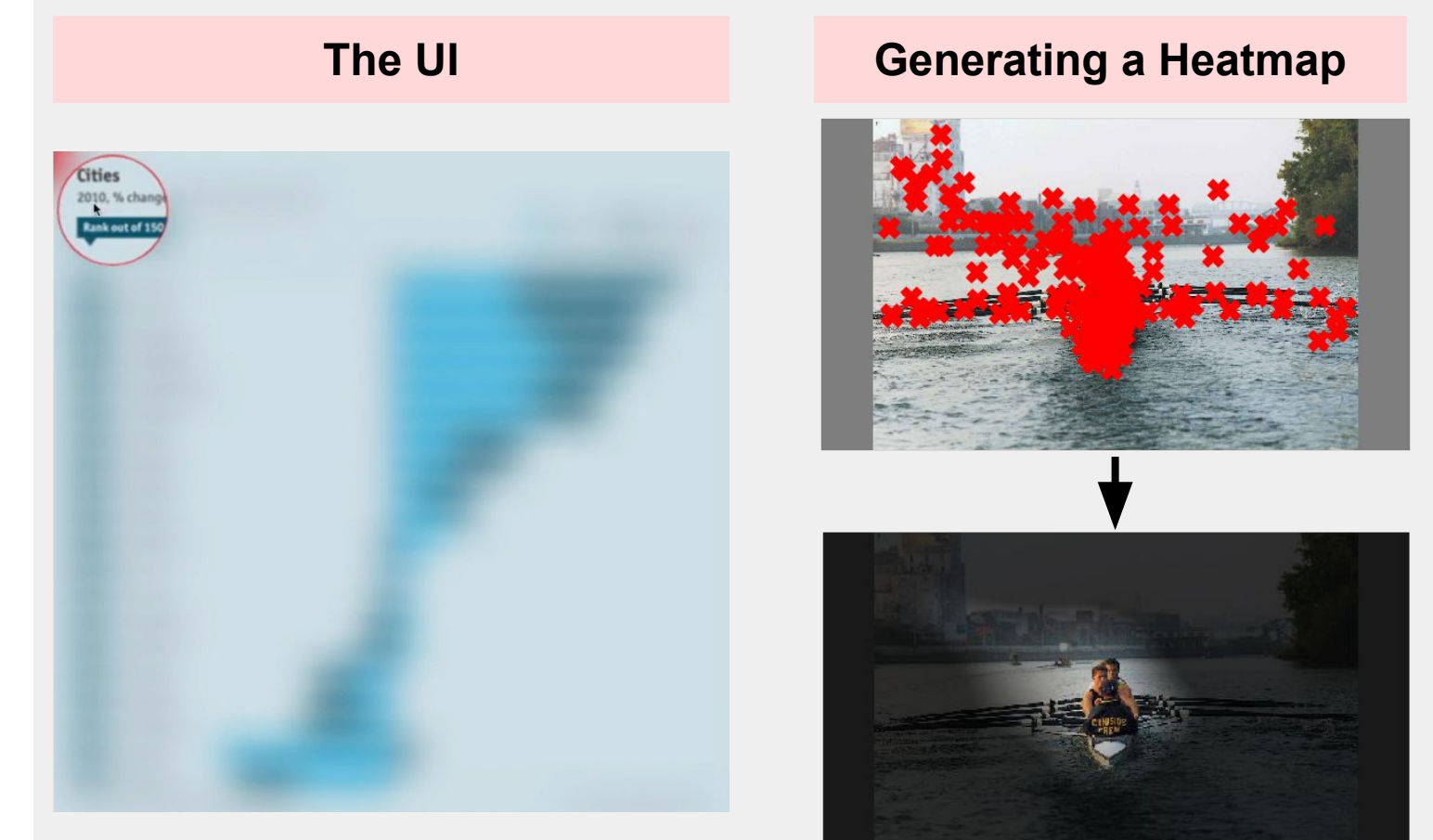
Which tool should I use?

We **ran experiments** with each of the interfaces in order to determine which **use cases** were appropriate for each.

	+	-
Zoom Maps	Works on multi-scale content, natural form of interaction	Coarse approximation of attention
Code Charts	Doesn't distort stimuli, approximates eye movements	Expensive, images must fit on screen
Import Annots	Produces clean segmentations, captures importance	Not ideal for natural images, importance > attention
Bubble View	Versatile, cheap	Distorts stimuli and timing

Image Type

Not all interfaces work with all image types.

ZoomMaps works great with multiscale images

ImportAnnots does not work as well with natural images

	Zoom Maps	Code Charts	Import Annots	Bubble View
Big / multiscale images	✓	✗	—	—
Natural images	✓	✓	—	✓
Non-natural images	✓	✓	✓	✓
Dynamic content	✗	✓	✗	✗

All interfaces work well on graphic designs and capture different aspects of attention

CodeCharts can work with video

| 500-5000 ms | 400 ms |

Cost

Cost depends on **how much data** we get per participant and **how much work** it is to use.

Cost per image

BubbleView	ZoomMaps	CodeCharts	ImportAnnots
$0.45	$0.72	$2.00	$3.00

- 15+ clicks/ participant
- 15 participants/ image for consistent data

- **Multiple** zoomed viewports per viewer
- **20** participants/ image

- **1** gaze point/ participant
- **50** participants/ image

- **Multiple** annotations/ participant
- **30** participants/ image
- **Extra work required** to segment elements
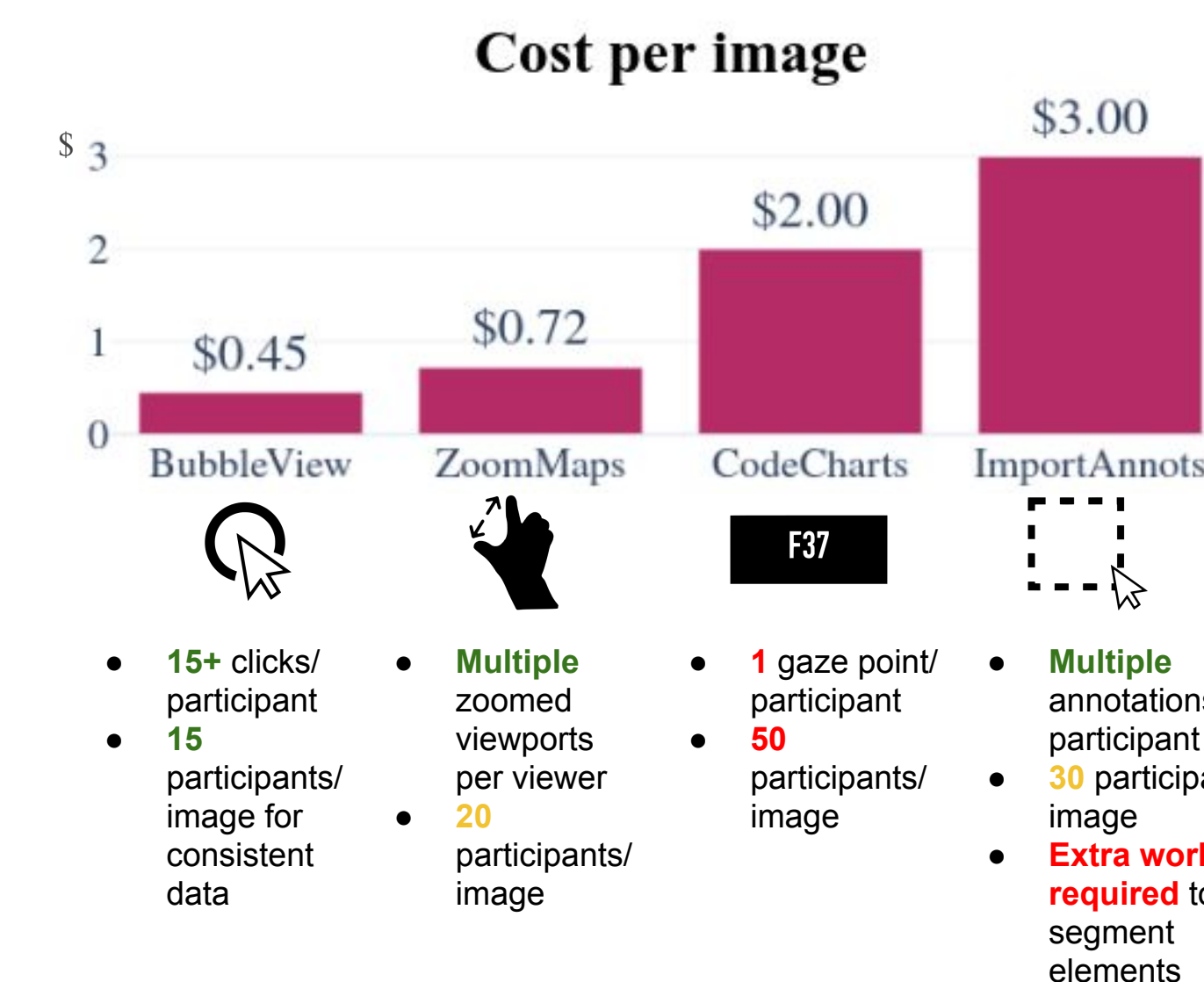
Similarity to Eye Movements

	Human Consistency	Code Charts	Bubble View	Zoom Maps	Import Annots
Similarity to eye movements*	0.86	**0.76**	0.62	0.59	0.51
% of Human Consistency	100%	**88%**	72%	69%	59%

Similarity to eye movements

| Eye Data | Code Charts | Bubble View | Zoom Maps | Import Annots |

*Similarity measured using Pearson's Correlation Coefficient comparing ground-truth to generated heatmaps

Saliency vs. Intentionality

We organize our interfaces on an "intentionality" scale based on the degree to which they measure **saliency (more spontaneous)** or **importance (more intentional)**.

Interfaces whose interaction method **requires more cognitive processing** are more intentional.

| Eye tracking | Code Charts | Bubble View | Zoom Maps | Import Annots |

Intentionality

Saliency → Importance