Anelise Newman
Barry McNamara
Camilo Fosco
Yun Bin Zhang

Introducing the TurkEyes Toolbox

UIs for crowdsourcing attention without an eye tracker

Pat Sukhum
Matthew Tancik
Nam Wook Kim
Zoya Bylinskii

Code, demos, and paper at
TurkEyes.mit.edu

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

CodeCharts *(self-report)*

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

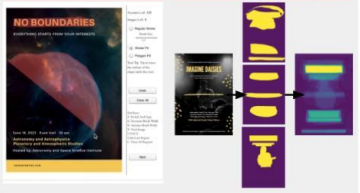
The UI

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

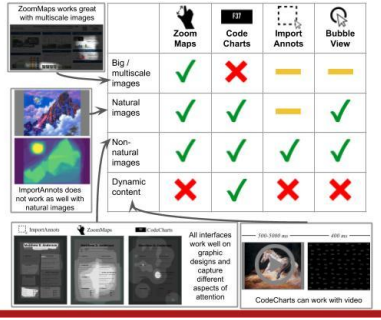
Which tool should I use?

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.



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

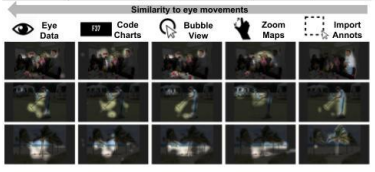
Cost

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

- **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%

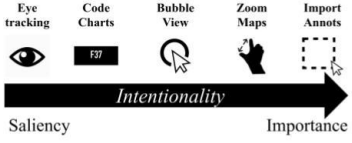


*Similarity measured using Pearson's Correlation Coefficient comparing ground-truth to generated heatmaps

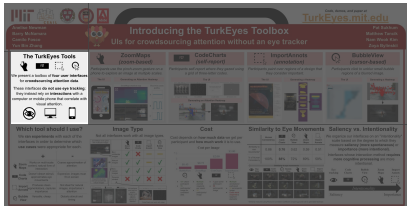
Saliency vs. Intentionality

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

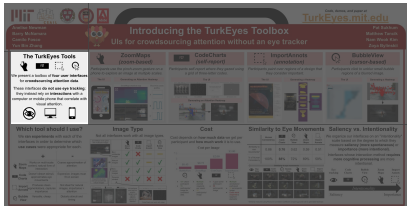
Interfaces whose interaction method **requires more cognitive processing** are more intentional.



The TurkEyes Tools



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

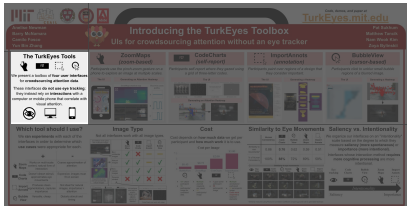


The TurkEyes Tools

F37

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

These interfaces **do not use eye tracking**;

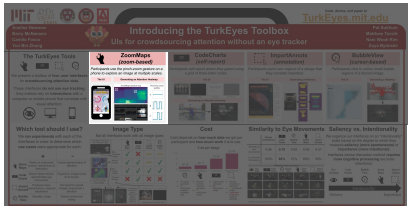


The TurkEyes Tools

F37

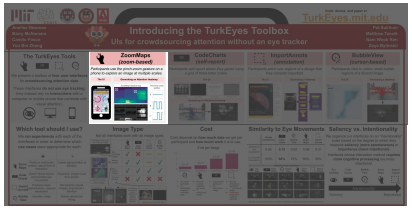
We present a toolbox of **four user interfaces** for **crowdsourcing attention data**.

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



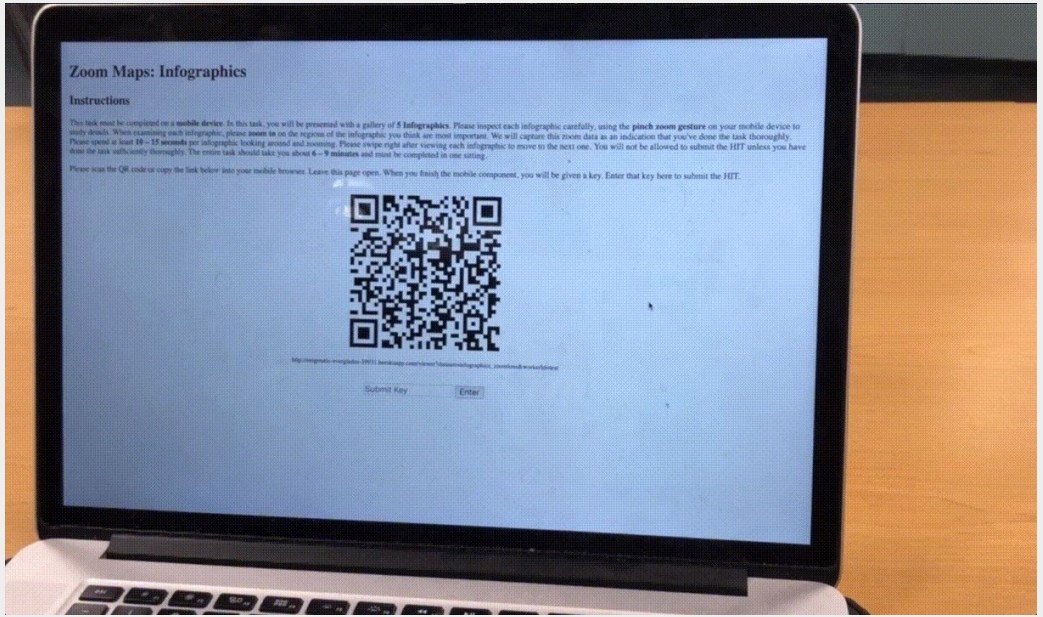
ZoomMaps *(zoom-based)*

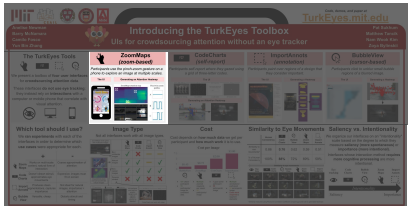
Participants use the pinch-zoom gesture on a phone to explore an image at multiple scales.



ZoomMaps *(zoom-based)*

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

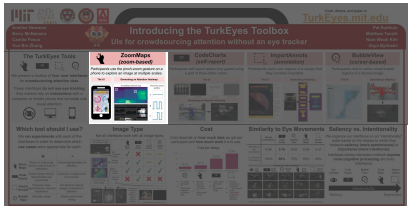




ZoomMaps *(zoom-based)*

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

The UI

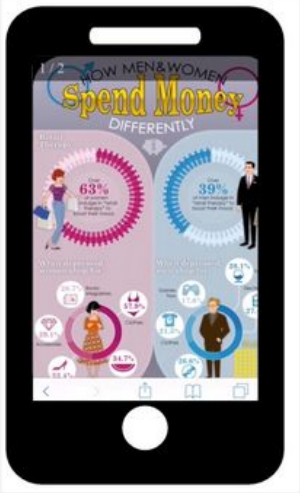


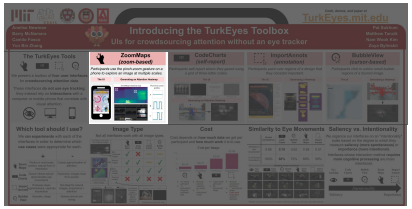
ZoomMaps *(zoom-based)*

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

The UI

Generating an Attention Heatmap

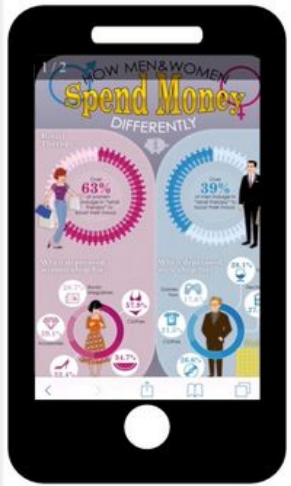




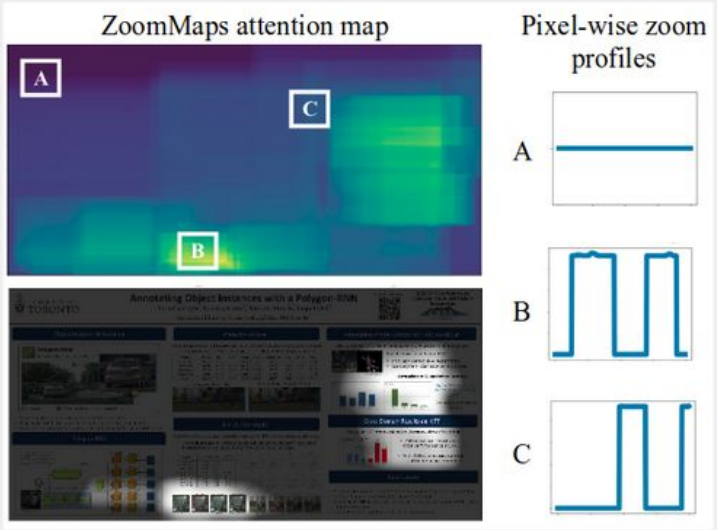
ZoomMaps *(zoom-based)*

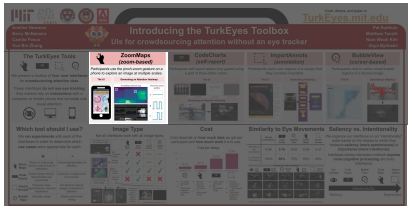
Participants use the pinch-zoom gesture on a phone to explore an image at multiple scales.

The UI



Generating an Attention Heatmap

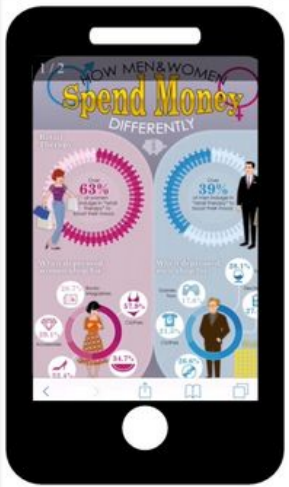




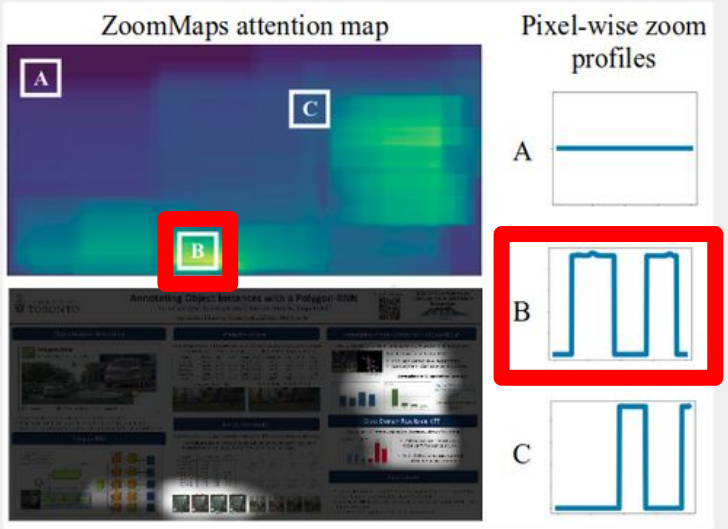
ZoomMaps *(zoom-based)*

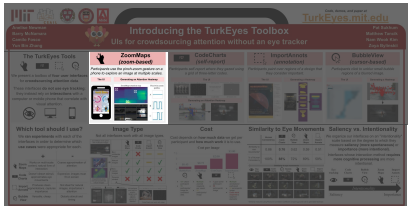
Participants use the pinch-zoom gesture on a phone to explore an image at multiple scales.

The UI



Generating an Attention Heatmap

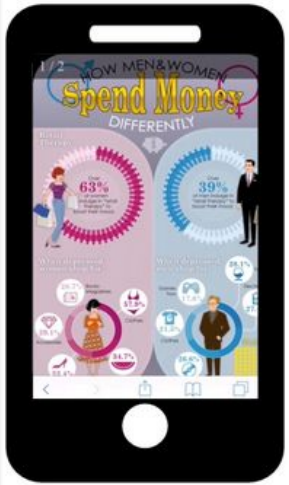




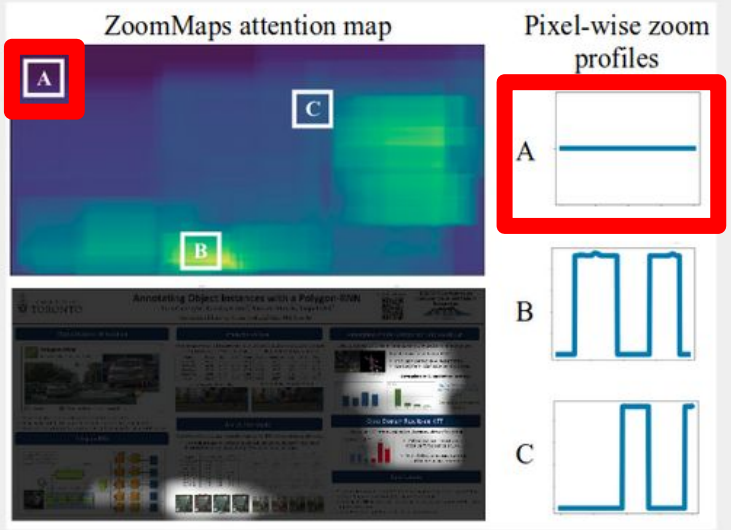
ZoomMaps *(zoom-based)*

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

The UI



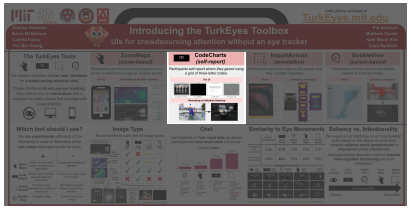
Generating an Attention Heatmap

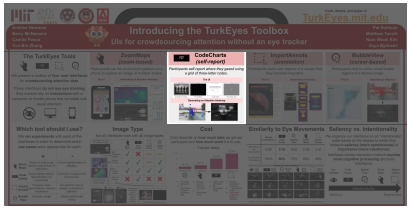


F37

CodeCharts *(self-report)*

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



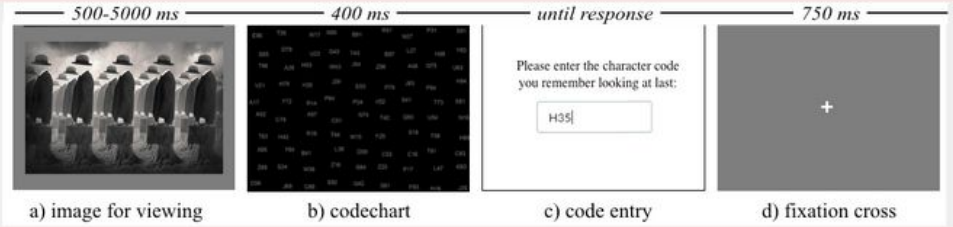


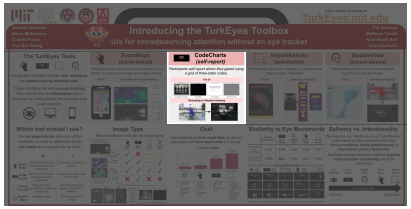
F37

CodeCharts *(self-report)*

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

The UI

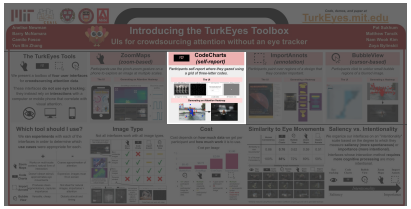




F37 CodeCharts *(self-report)*

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

Image Attention

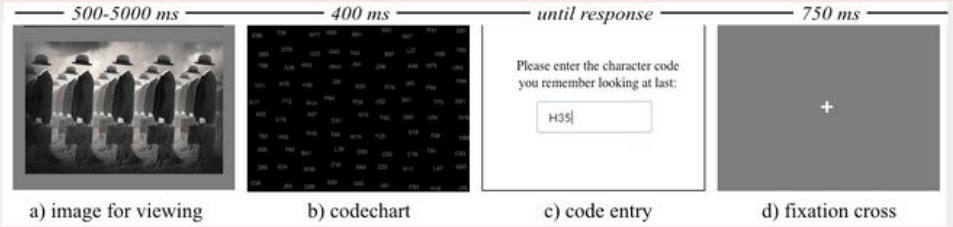


F37

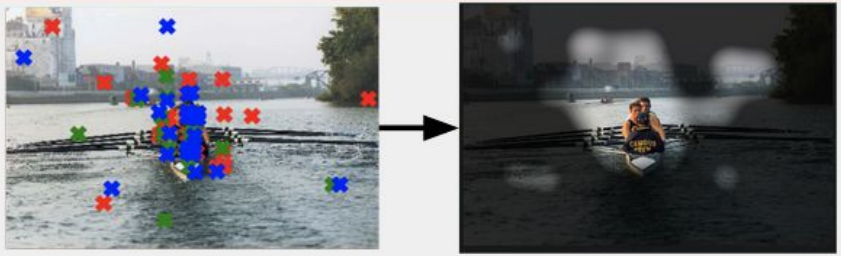
CodeCharts *(self-report)*

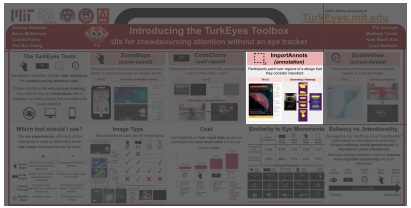
Participants self-report where they gazed using a grid of three-letter codes.

The UI



Generating an Attention Heatmap



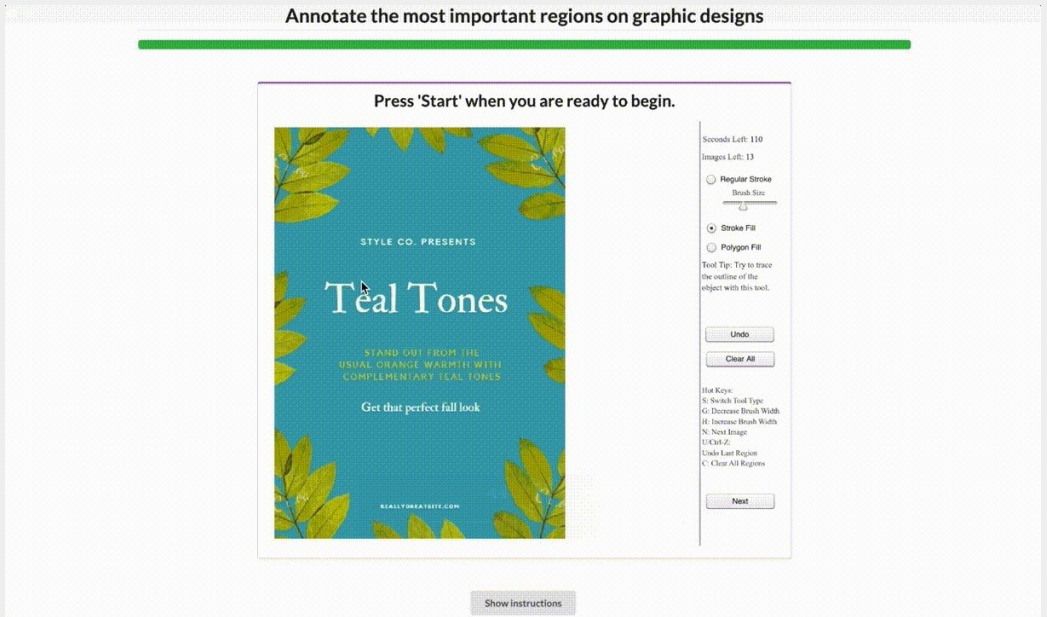


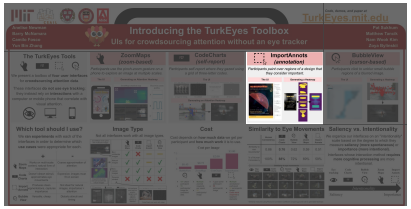
ImportAnnots *(annotation)*

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

ImportAnnots *(annotation)*

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

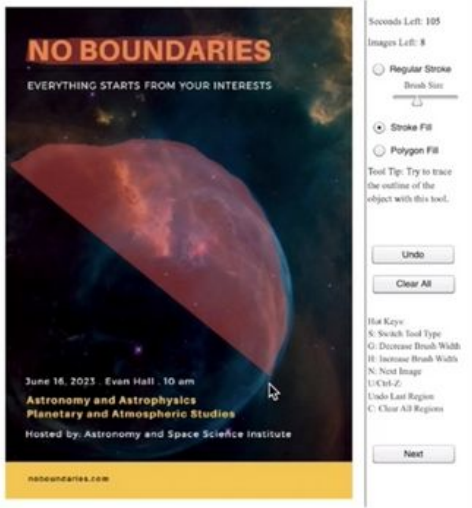




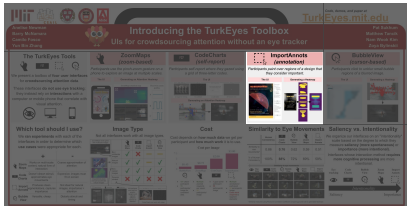
ImportAnnots *(annotation)*

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

The UI



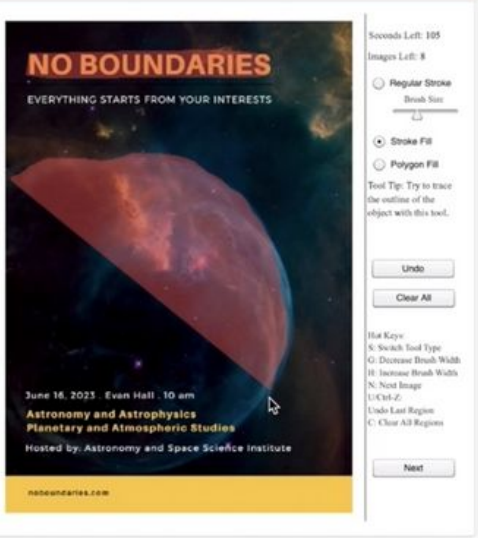
Generating a Heatmap



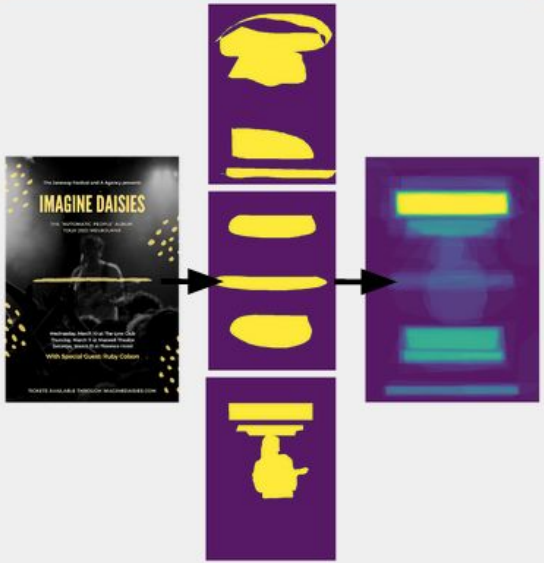
ImportAnnots *(annotation)*

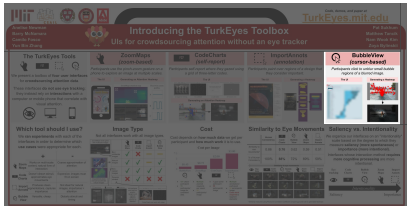
Participants paint over regions of a design that they consider important.

The UI



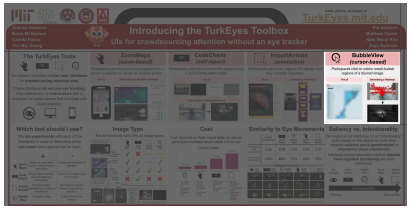
Generating a Heatmap





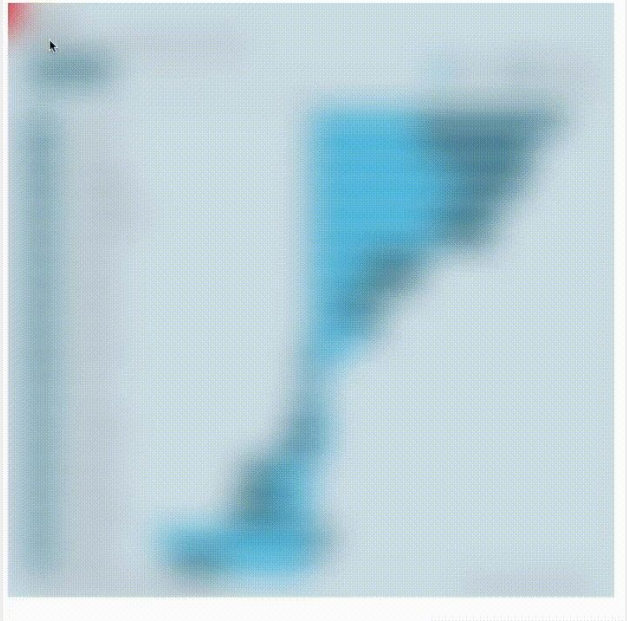
BubbleView *(cursor-based)*

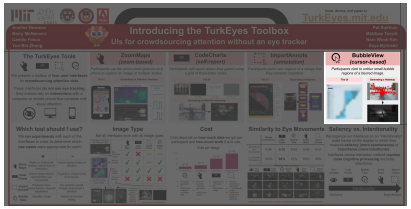
Participants click to unblur small bubble regions of a blurred image.



BubbleView *(cursor-based)*

Participants click to unblur small bubble regions of a blurred image.



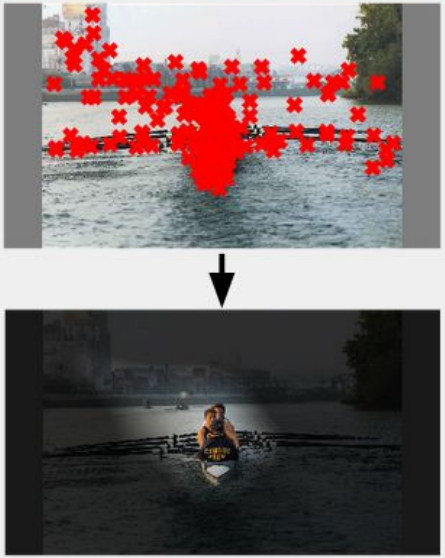


BubbleView *(cursor-based)*

Participants click to unblur small bubble regions of a blurred image.

The UI

Generating a Heatmap



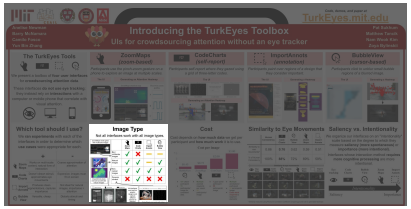
Which tool should I use?

Which tool should I use?

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

Image Type

Not all interfaces work with all image types.



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

Image Type

Not all interfaces work with all image types.

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

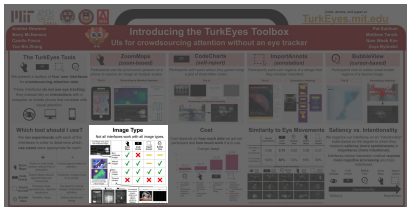
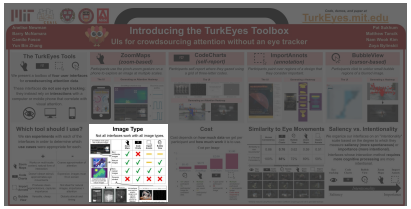
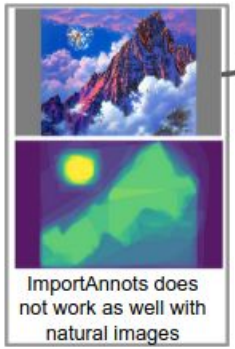


Image Type

Not all interfaces work with all image types.

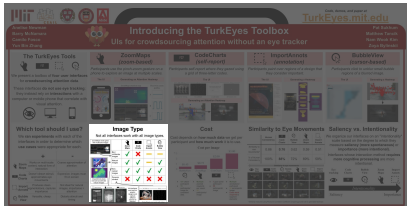




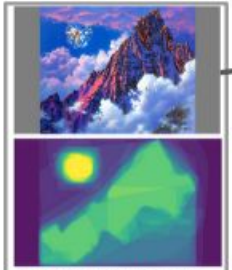
	Zoom Maps	Code Charts	Import Annots	Bubble View
Big / multiscale images	✓	✗	–	–
Natural images	✓	✓	–	✓
Non-natural images	✓	✓	✓	✓
Dynamic content	✗	✓	✗	✗

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	✗	✓	✗	✗

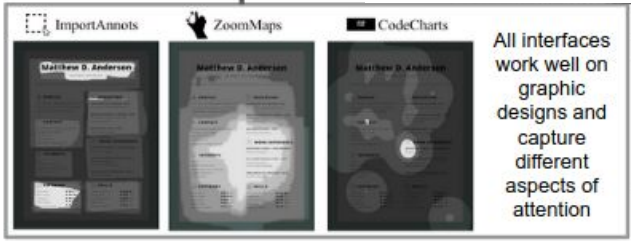
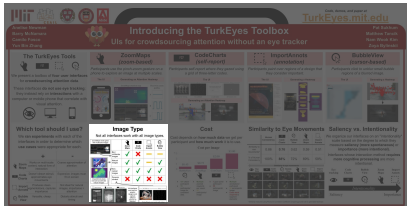
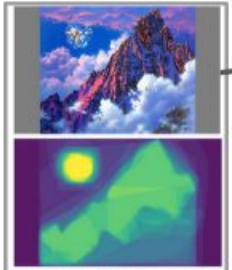


Image Type

Not all interfaces work with all image types.

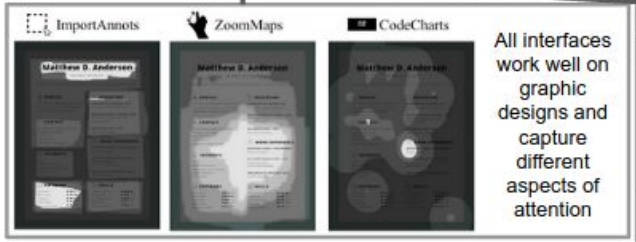


ZoomMaps works great with multiscale images

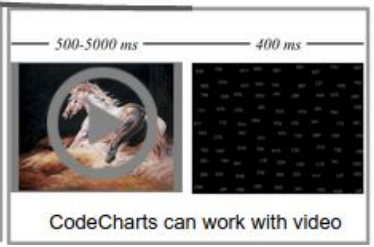


ImportAnnots does not work as well with natural images

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



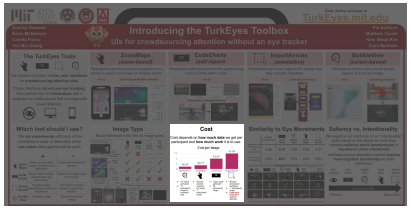
All interfaces work well on graphic designs and capture different aspects of attention



CodeCharts can work with video

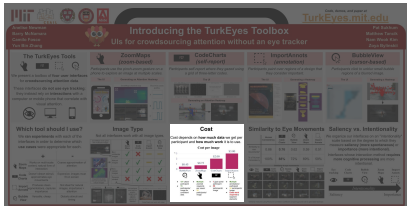
Cost

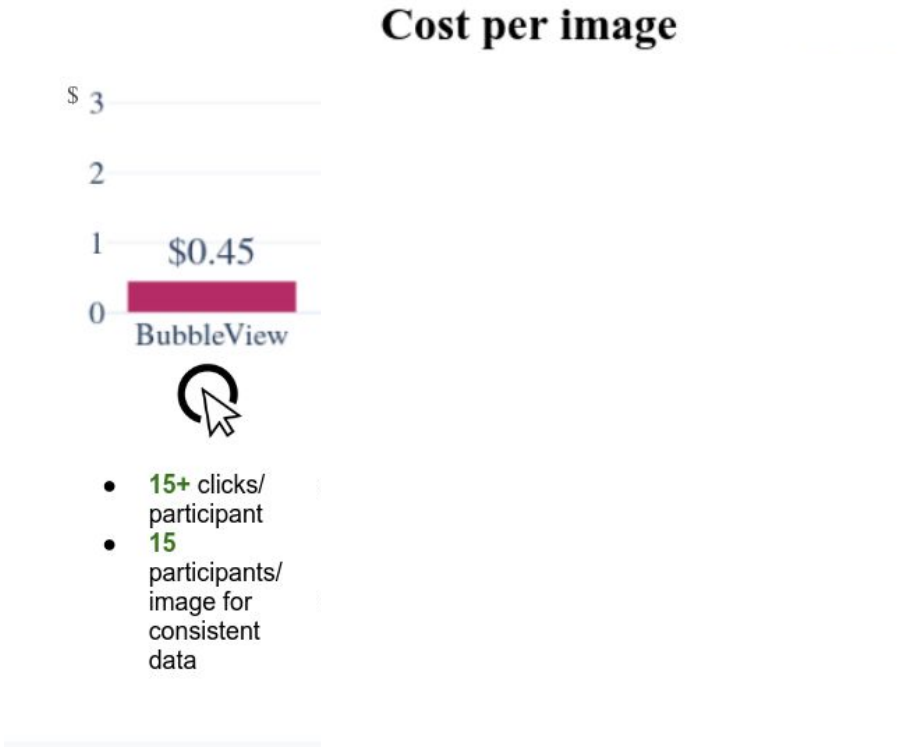
Cost depends on **how much data** we get per participant and **how much work** it is to use.



Cost

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

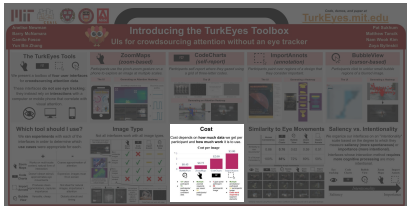


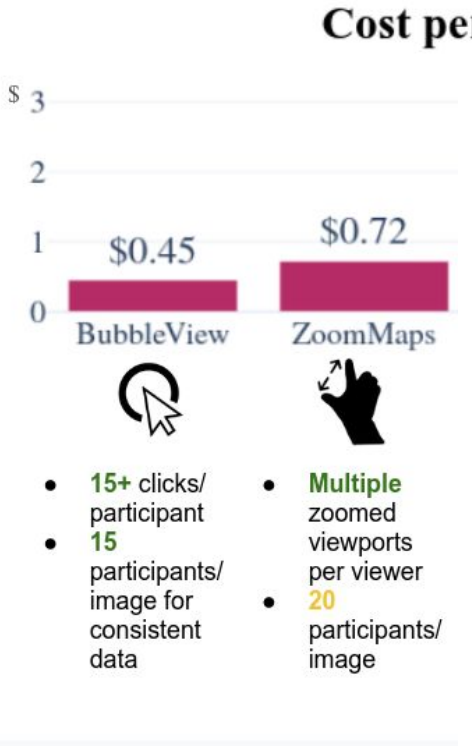


- **15+** clicks/ participant
- **15** participants/ image for consistent data

Cost

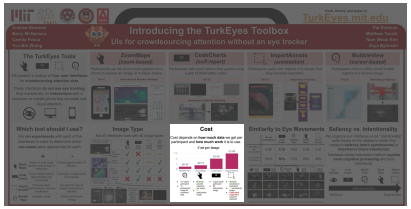
Cost depends on **how much data** we get per participant and **how much work** it is to use.

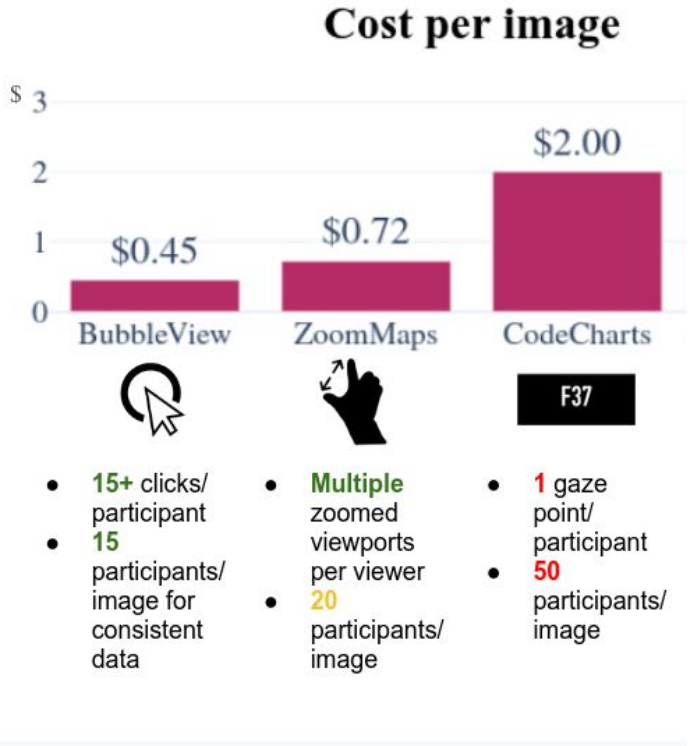




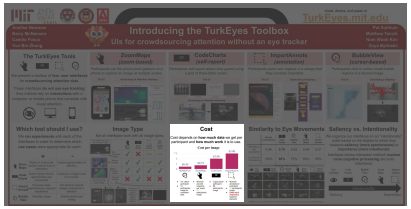
Cost

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

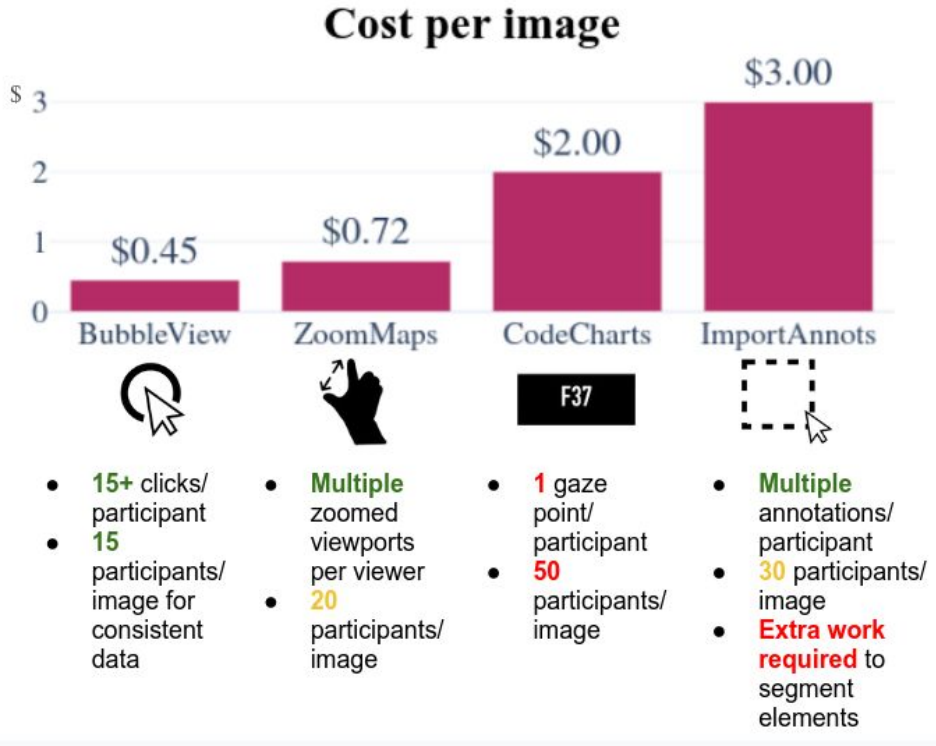




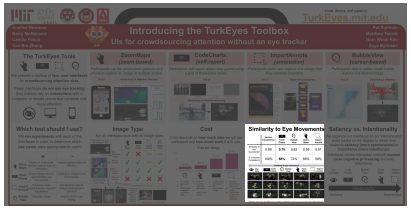
Cost



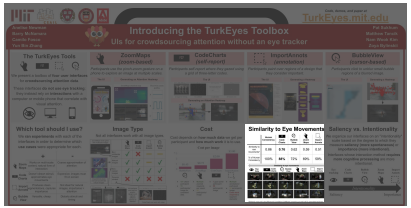
Cost depends on **how much data** we get per participant and **how much work** it is to use.



Similarity to Eye Movements



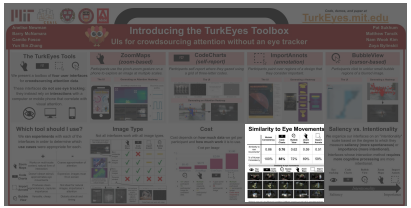
Similarity to Eye Movements



	Human Consistency
Similarity to eye movements*	0.86
% of Human Consistency	100%

Similarity to eye movements

Similarity to Eye Movements



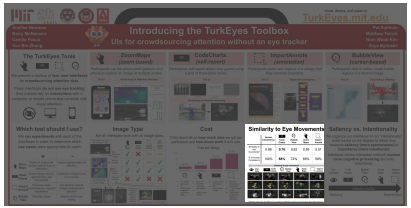
	Human Consistency	F37 Code Charts
Similarity to eye movements*	0.86	**0.76**
% of Human Consistency	100%	**88%**

Similarity to eye movements

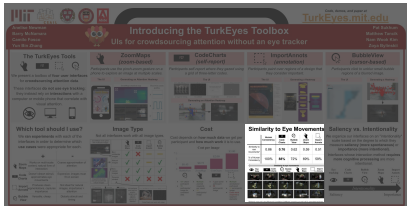
Similarity to Eye Movements

	Human Consistency	F37 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%

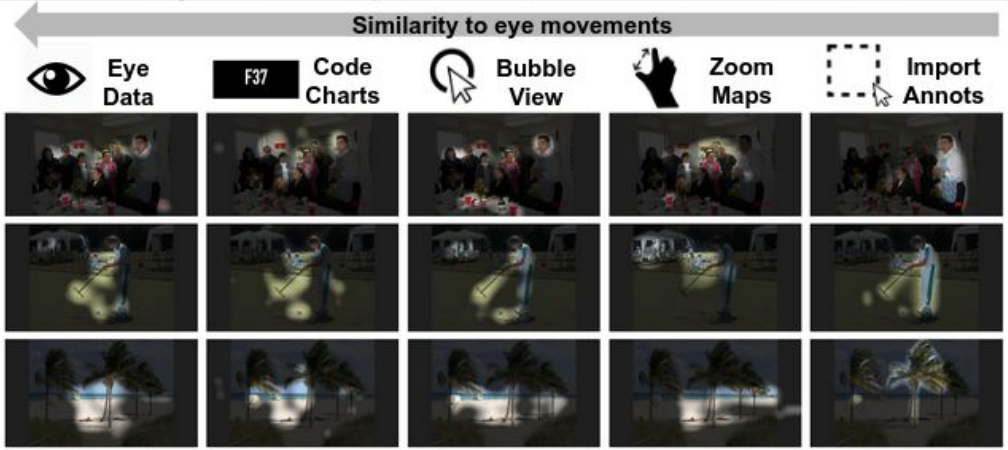
Similarity to eye movements



Similarity to Eye Movements

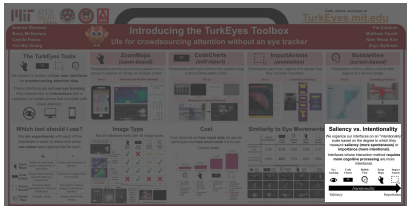


	Human Consistency	F37 Code Charts	Bubble View	Zoom Maps	Import Annots
Similarity to eye movements*	0.86	**0.76**	0.62	0.59	0.51
% of Human Consistency	100%	**88%**	72%	69%	59%



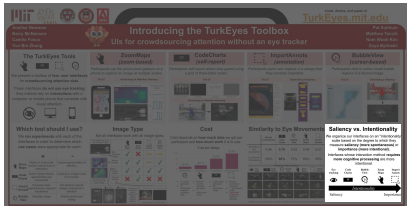
*Similarity measured using Pearson's Correlation Coefficient comparing ground-truth to generated heatmaps

Saliency vs. Intentionality



Saliency vs. Intentionality

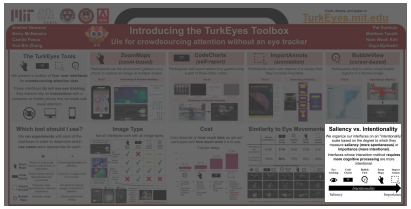
We organize our interfaces on an "intentionality" scale based on the degree to which they measure **saliency (more spontaneous)** or **importance (more intentional).**



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).**

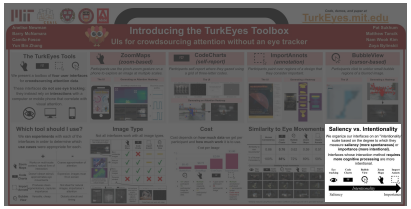
Interfaces whose interaction method r**equires more cognitive processing** are more intentional.



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).**

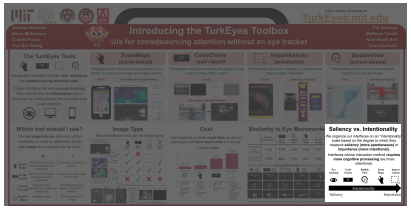
Interfaces whose interaction method r**equires more cognitive processing** are more intentional.

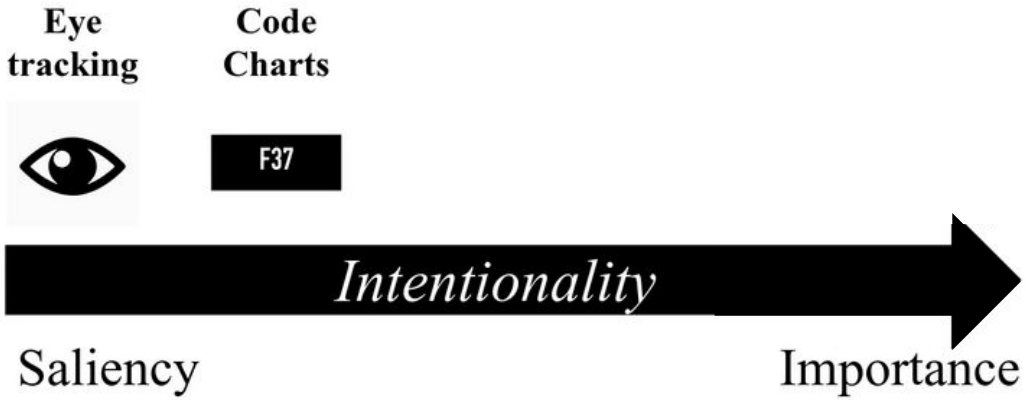


Saliency vs. Intentionality

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 r**equires more cognitive processing** are more intentional.

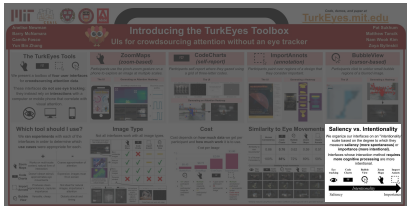


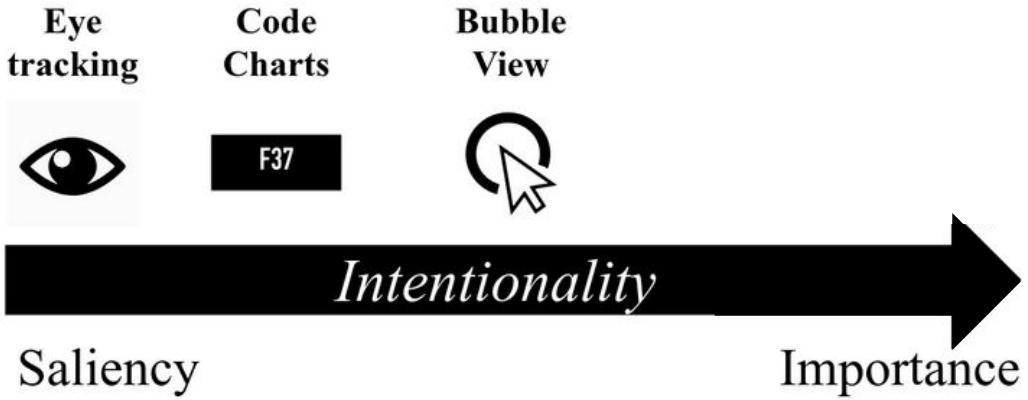


Saliency vs. Intentionality

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 r**equires more cognitive processing** are more intentional.

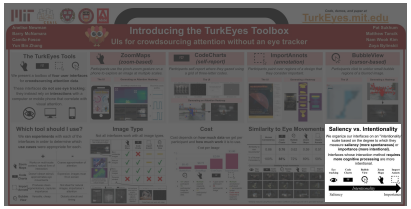


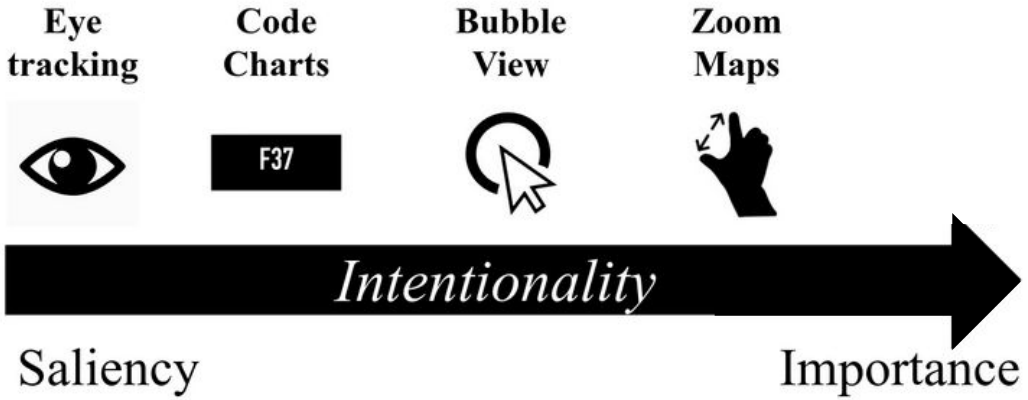


Saliency vs. Intentionality

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 r**equires more cognitive processing** are more intentional.

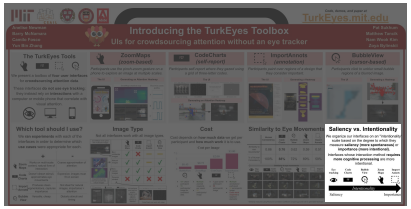


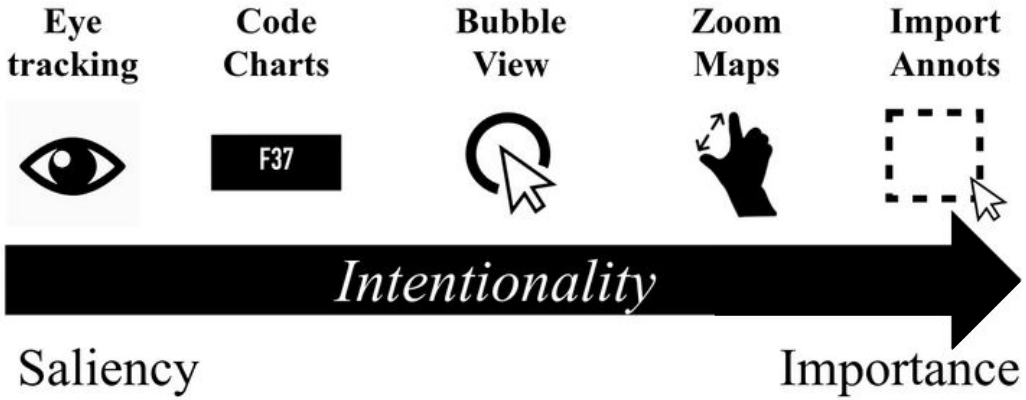


Saliency vs. Intentionality

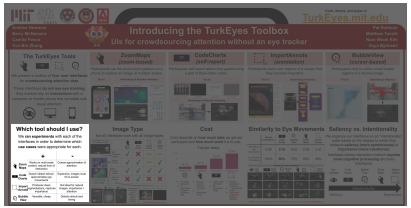
We organize our interfaces on an "intentionality" scale based on the degree to which they measure **saliency (more spontaneous)** or **importance (more intentional).**

Interfaces whose interaction method r**equires more cognitive processing** are more intentional.

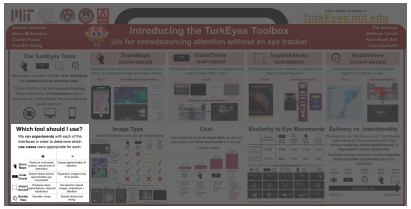




Which tool should I use?

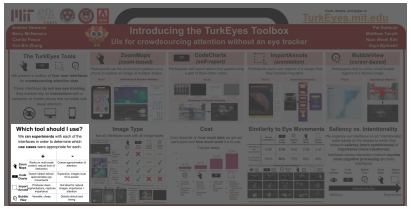


Which tool should I use?



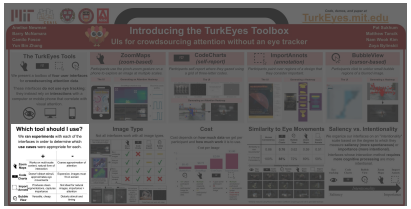
	+	-
Zoom Maps	Works on multi-scale content, natural form of interaction	Coarse approximation of attention

Which tool should I use?



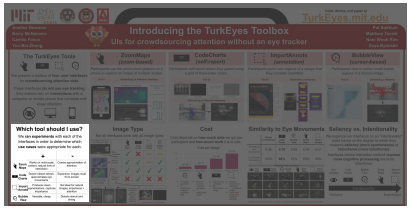
	+	-
Zoom Maps	Works on multi-scale content, natural form of interaction	Coarse approximation of attention
F37 **Code Charts**	Doesn't distort stimuli, approximates eye movements	Expensive, images must fit on screen

Which tool should I use?



	+	-
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

Which tool should I use?



	+	-
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

Code, demos, and paper at

TurkEyes.mit.edu

Anelise Newman
Barry McNamara
Camilo Fosco
Yun Bin Zhang

Introducing the TurkEyes Toolbox

UIs for crowdsourcing attention without an eye tracker

Pat Sukhum
Matthew Tancik
Nam Wook Kim
Zoya Bylinskii

Code, demos, and paper at
TurkEyes.mit.edu

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

CodeCharts *(self-report)*

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

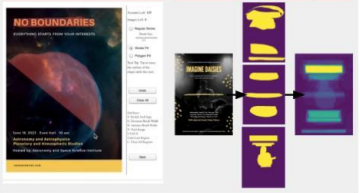
The UI

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

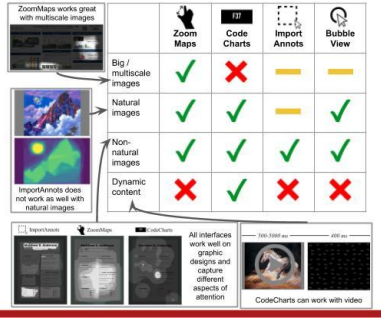
Which tool should I use?

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.



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

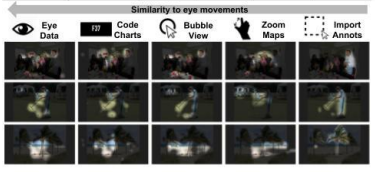
Cost

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

- **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 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.