Supplementary Materials for "VisRecall++: Analysing and Predicting Recallability of Information Visualisations from Gaze Behaviour"

YAO WANG*, University of Stuttgart, Germany
YUE JIANG, Aalto University, Finland
ZHIMING HU, University of Stuttgart, Germany
CONSTANTIN RUHDORFER, University of Stuttgart, Germany
MIHAI BÂCE†, KU Leuven, Belgium
ANDREAS BULLING, University of Stuttgart, Germany

ACM Reference Format:
Yao Wang, Yue Jiang, Zhiming Hu, Constantin Ruhdorfer, Mihai Bâce, and Andreas Bulling. 2024. Supplementary Materials for "VisRecall++: Analysing and Predicting Recallability of Information Visualisations from Gaze Behaviour". 1, 1 (April 2024), 8 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

This document contains the explanatory texts for experiment files (Section 1) and gaze data processing procedures (Section 2). We include images illustrating the recording setup and screenshots during the encoding phase and recall phase of the study (Figure 4), examples of visualisations produced by VisRecall++ with AOI annotations (Figure 5) and with fixation contours (Bell Curve), a scarf plot, and tables displaying percentage fixation duration (Figure 6). Additionally, we present statistics for the high and low recallability groups in Table 1. These visualisations and statistics are included as supplements to the main manuscript.

1 EXPERIMENT FILES

Since the EyeLink 1000 Plus eye tracker generates one single file (.edf) as the raw gaze output and one device session log, we recorded several supportive files to facilitate later data processing. Experimental files are:

Raw Gaze Output. This file contains recorded raw gaze data in the encoding and recording phase of twenty visualisations. We cut this file into several viewing periods for every stimulus.

*Corresponding author
†A significant part of this work was conducted while at the University of Stuttgart

Authors' addresses: Yao Wang, University of Stuttgart, Stuttgart, Germany, yao.wang@vis.uni-stuttgart.de; Yue Jiang, Aalto University, Espoo, Finland, yue.jiang@aalto.fi; Zhiming Hu, University of Stuttgart, Stuttgart, Germany, zhiming.hu@vis.uni-stuttgart.de; Constantin Ruhdorfer, University of Stuttgart, Stuttgart, Germany, constantin.ruhdorfer@vis.uni-stuttgart.de; Mihai Bâce, KU Leuven, Leuven, Belgium, mihai.bace@kuleuven.be; Andreas Bulling, University of Stuttgart, Stuttgart, Germany, andreas.bulling@vis.uni-stuttgart.de.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2024 Association for Computing Machinery.
XXXX-XXXX/2024/4-ART $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

Device Session Log. This log file is created by Weblink, a software for linking the eye tracker and the study web applications. It records events on the computer, such as mouse clicks and scroll events. Log events here are also measured in milliseconds, but relative to the start of the Weblink software.

Experiment Group. This is a file describing which information visualisations were presented and in which order and which questions were asked in between.

Subject Fixations. This file contains all eye tracking events of the subject, i.e. fixations, saccades, but also eye validation scores etc. These events are logged in milliseconds counted from the device's startup.

Study Application Log. This log file contains which events occurred in the study application. Such events include: pressing the next button and submitting the study results. Here events are logged in standard Unix timestamps.

AOI Annotations. These polygons describe the areas of interest (AOIs) for visualisations. Ten types of AOIs were annotated (see main manuscript for definition). Every visualisation has a single LabelMe [1] formatted .txt file.

Gazealytics Supportive Files. These are the files to support the usage of Gazealytics[2], a web-based visual eye tracking analytics toolkit. The toolkit supports a combination of AOI and non-AOI comparative analysis and filter options to visually explore gaze patterns across time and space. A single .zip is required to conduct the analysis in this toolbox for every visualisation. We provide a conversion script to convert LabelMe formatted annotation files to the AOIs.json required by Gazealytics.

2 GAZE DATA PROCESSING

The files described in Section 1 provide timestamps in different time systems. To retrieve scanpaths from raw gaze data, aligning and matching these timestamps is crucial. Primarily, since the fixation events are in another timestamp format from the application events, we perform a proper alignment on which visualisations were presented at which time. A mapping from the web page's domain onto the image domain is necessary. Apart from aligning timestamps, we must compare the application logs against the study description to understand which button click corresponds to which displayed visualisation. Lastly, we are mainly interested in the gaze behaviour of subjects on the images rather than their behaviour on the complete web application (which is mostly empty and void of distractions).

2.1 Timestamp Matching and Alignment.

As previously noted, accurate timestamp matching and aligning is paramount. For this, consider how the previously mentioned log files are related. First, the application log gives us the actual Unix time, which we want the other timestamps to convert. Since the application log only logs internal program states, we find recognisable events across the log files to allow for matching (compared with Listing 1).

For matching, we exploit the following observation: the time someone enters the recall stage is known to us (here 1638374473255). This stage is only entered after a user scrolls down on the first page and clicks the "Next" button. We can deduce that there has to be a mouse click event at the very start, which occurs precisely after a mouse scroll event.

[1] https://github.com/CSAILVision/LabelMeAnnotationTool
[2] https://github.com/gazealytics/gazealytics-master

```
---- Session started: Wed Dec 01 2021 17:00:53 GMT+0100 ----
Entering recall stage: Wed Dec 01 2021 17:01:00 GMT+0100
1638374460512
next button clicked for task:0, at: Wed Dec 01 2021 17:01:13 GMT+0100
1638374473255
...
```

Fig. 1. Excerpt from application log

```
...
58. 14363000 Main Message upon Start: 7701143 Web Component
...
109 14369647 Main EyeLinkController Message - MouseClick Left 1351 944
110 14369647 Main EyeLinkController Message - MouseUp Left 1351 944
111 14369659 Web Incoming 21 Scroll Event: [0 0] Viewport size: 1920x1080
...
```

Fig. 2. Simplified excerpt from device session log

```
...
MSG 7701680 _MousePosition_ 1921 0
EFIX L 7701115 7701685 570 989.3 537.4 2852
SSACC L 7701686
EFIX R 7701115 7701688 573 961.0 526.5 2799
SSACC R 7701689
...
```

Fig. 3. Excerpt from subject fixations file containing two fixation events.

The relevant mouse click event is on line 110 in the example Listing 2. Next, one must match the device session log to the fixation events. Fixations are in the format *EFIX <eye> <start_time> <end_time> <dur> <axp> <ayp> <aps>*. See an example in Listing 2.

Since the device session log mentions the device time on line 58 in the given example, this mapping can be easily read out. We thus construct a tuple that gives the total relationship between these, i.e. given

$$t_{\text{application}} \leftrightarrow t_{\text{device_session}} \tag{1}$$

and

$$t_{\text{main_message}} \leftrightarrow t_{\text{eyetreacker_time}} \tag{2}$$

we construct

$$(t_{\text{application}}, t_{\text{device_session}}, t_{\text{main_message}}, t_{\text{eyetreacker_time}}) \tag{3}$$

. With this mapping, one can obtain the Unix time of an eye fixation event by calculating the following:

$$t_{\text{offset}} = t_{\text{application}} - (t_{\text{eyetreacker_time}} + t_{\text{device_session}} - t_{\text{main_message}}). \quad (4)$$

For instance, with the given example fixation EFIX L 7701115 ... we calculate $t_{\text{offset}} = (1638374460512 - (7701143 + 14369647 - 14363000))$ and obtain $7701115 + t_{\text{offset}} = 1638374453837$. We repeat this process for every experiment run.

2.2 Eye Fixation Mapping

Fixations on the screen need to be mapped onto the relevant visualisation image domain. Furthermore, since all images displayed are fitted and stretched into a 1066×800 box such that one of the image axes fits the box axis, these fixations must be correctly translated to the original image sizes. To this end, we calculate the fitted image size and then filter and re-stretch the coordinates of the filtered fixations.

To increase the accuracy of our performed analysis, we determine the eye with better validation accuracy as determined by the SR Research[3] validation accuracy and focus our analysis on this eye. We do so for every subject individually. All our subjects show the best possible class of validation scores ("Good") on the kept eye.

2.3 Output Format

Our produced output is consistent with the published version of the MASSVIS dataset[4].

Listing 1. Schematic display of the output structure where {stage} denotes enc or recall.

```
eyetracking/csv_files/
|    +-- fixationsByVis/
|    |    +-- {img_name}/
|    |    |    +-- {stage}/
|    |    |    |    +-- {Subject}.csv
|    +-- strings/
|    |    +-- {Subject}/
|    |    |    +-- {stage}/
|    |    |    |    +-- {img_name}.txt
```

We display a schematic view of the output in Listing 1. The authors hope this correspondence makes pickup and use of the dataset easier, increasing adoption in other works. The fixation $*$.csv files have columns <index>, <x>, <y>, <duration> also consistent with previous work. The strings folder contains files containing scanpath descriptions on visual elements where each $*$.txt file denotes a single scanpath over visual elements on the image for the subject.

2.4 Timestamp Matching Validation

Since eye fixations must be precise in time, we further describe how we have validated our previous processing pipeline and results. We validate the previously described timestamp tuples to avoid displacements of fixations in timestamp matching described in Section 2.1. We construct another tuple for each timestamp-matching tuple that consists of the later timestamp in the device session and application log. To this end, we find another mouse click event that corresponds to an entry in

[3] https://www.sr-research.com/support/attachment.php?aid=1376
[4] https://github.com/massvis/eyetracking

Table 1. Statistical results of fixation durations and saccade amplitudes of high and low recallability groups in 10 trials. Stars indicate the statistical significance of the difference between scanpaths with high and low recallability (*: p < .05; **: p < .01; ***: p < .001).

Trial	Threshold	Participants	Fixation Duration	Saccade Amplitude
T1-high	0.493	11	244.65 (145.16)	147.62 (420.80)
T1-low		6	238.88 (133.08)	138.56 (413.53)
T2-high	0.435	9	210.96 (124.19)	235.74 (472.22)
T2-low		8	219.96 (118.36)***	221.00 (419.40)
T3-high	0.387	8	218.09 (130.99)	123.28 (259.41)
T3-low		7	219.96 (118.36)	122.27 (247.49)
T4-high	0.471	9	216.44 (126.77)*	149.82 (321.47)*
T4-low		9	210.93 (140.78)	136.96 (300.56)
T5-high	0.431	9	200.14 (117.93)	213.79 (475.02)***
T5-low		7	237.85 (121.56)***	225.31 (503.34)
T6-high	0.403	8	212.72 (122.29)	183.37 (438.20)
T6-low		9	221.46 (134.98)***	171.51 (399.08)
T7-high	0.432	12	218.10 (124.14)***	122.09 (211.42)
T7-low		5	208.90 (138.32)	130.88 (225.27)
T8-high	0.454	10	212.99 (126.07)	210.09 (398.95)
T8-low		8	230.59 (138.52)***	203.67 (399.98)
T9-high	0.408	9	228.84 (123.12)	190.72 (488.59)***
T9-low		8	238.47 (142.19)***	153.99 (379.14)
T10-high	0.407	12	219.93 (129.31)***	325.91 (856.30)
T10-low		6	211.31 (180.24)	311.11 (843.08)

the application log and calculate the difference with the original timestamps:

$$\delta = t_{\text{application}} - t_{\text{device_session}} \tag{5}$$

, as well as the newly generated ones:

$$\delta_{\text{val}} = t_{\text{application_val}} - t_{\text{device_session_val}} \tag{6}$$

. Since the calculated conversion holds for all timestamps in the device session log, we expect the difference to be identical

$$\delta - \delta_{\text{val}} \stackrel{!}{=} 0 \tag{7}$$

. For practical reasons, and since the resolution of these timestamps is one millisecond, we allow for a difference of up to 1 millisecond to deal with rounding errors and the like. We exclude 6 out of 160 total experiment runs based on this validation step.

Fig. 4. Left: Recording setup consisting of a high-speed stationary eye tracker, a high-resolution display, and a chin rest. Top: a screenshot of the encoding stage of our study. Bottom: a screenshot of the recall stage of our study.

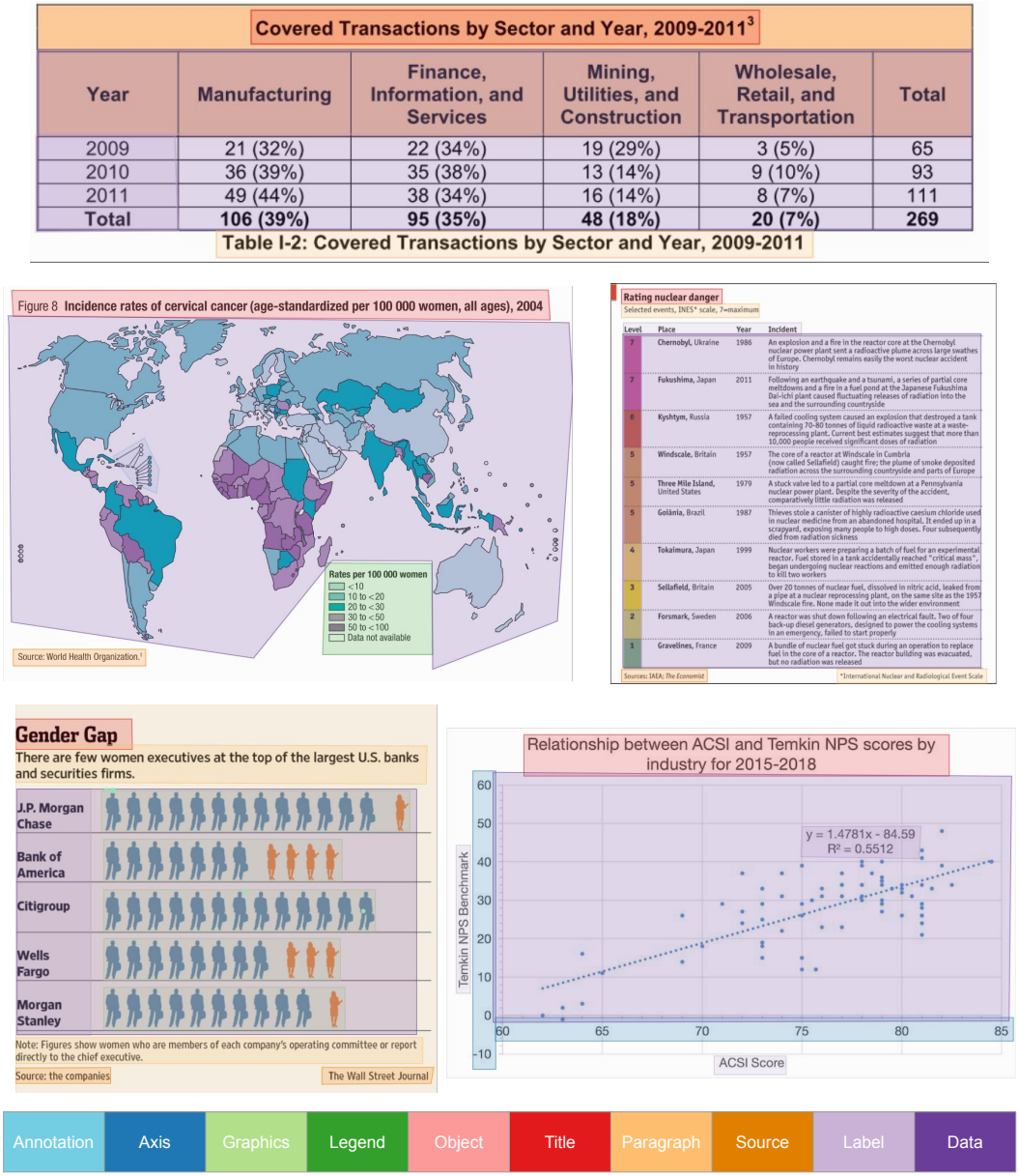
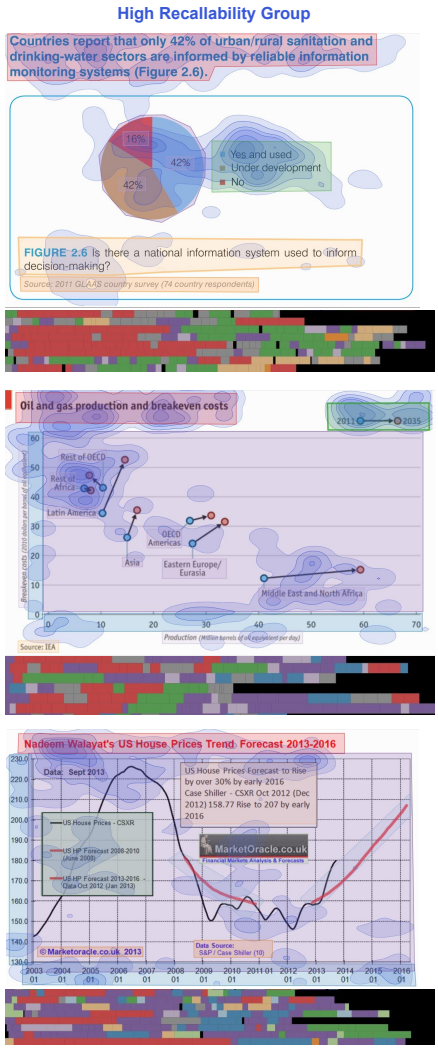


Fig. 5. Five samples of annotated AOIs overlaid on the same visualisation from VisRecall++.



	High	Low
Legend	**26.1**	20.5
Title	33.8	45.4
Paragraph	7.8	3.7
Source	1.1	0
Label	3.35	2.15
Data	14.4	**16.5**

	High	Low
Axis	**7.7**	5.1
Legend	**13.4**	1.8
Title	26.0	21.7
Paragraph	0	0.4
Label	8.1	9.0
Data	38.6	**57.1**

	High	Low
Annotation	1.96	2.11
Axis	**13.5**	4.2
Legend	8.95	16.4
Title	20.7	27.6
Paragraph	4.5	8.6
Source	0.3	2.1
Label	2.9	3.6
Data	67.5	67.3

Fig. 6. Three examples from VisRecall++, each divided into groups of high and low recallability group. The visualisations included fixation contours (Bell Curve), a scarf plot, and tables displaying percentage fixation duration. The high recallability group usually has longer timeline stripes, lower percentage dwell time on Data (D), and longer percentage dwell time on Axes (X) and Legends (L).