FakeNewsPerception: an eye movement dataset on the perceived believability of news stories

Ömer Sümer, Efe Bozkir, Thomas Kübler, Sven Grüner, Sonja Utz, Enkelejda Kasneci

Angaben zur Veröffentlichung / Publication details:

Sümer, Ömer, Efe Bozkir, Thomas Kübler, Sven Grüner, Sonja Utz, and Enkelejda Kasneci. 2021. "FakeNewsPerception: an eye movement dataset on the perceived believability of news stories." *Data in Brief* 35: 106909. https://doi.org/10.1016/j.dib.2021.106909.

Data in Brief 35 (2021) 106909

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

FakeNewsPerception: An eye movement dataset on the perceived believability of news stories

Ömer Sümer[a], Efe Bozkir[a], Thomas Kübler[a], Sven Grüner[b], Sonja Utz[c,d], Enkelejda Kasneci[a,*]

[a] Human-Computer Interaction, Department of Computer Science, University of Tübingen, Tübingen, Germany
[b] Faculty of Natural Sciences, Martin Luther University Halle-Wittenberg, Halle, Germany
[c] Leibniz-Institut für Wissensmedien, Tübingen, Germany
[d] Department of Psychology, University of Tübingen, Tübingen, Germany

ARTICLE INFO

Article history:
Received 10 December 2020
Revised 17 February 2021
Accepted 22 February 2021
Available online 1 March 2021

Keywords:
Fake news perception
Eye tracking dataset
Scanpath comparison
Eye movements
Reading comprehension
Misinformation

ABSTRACT

Extensive use of the internet has enabled easy access to many different sources, such as news and social media. Content shared on the internet cannot be fully fact-checked and, as a result, misinformation can spread in a fast and easy way. Recently, psychologists and economists have shown in many experiments that prior beliefs, knowledge, and the willingness to think deliberately are important determinants to explain who falls for fake news. Many of these studies only rely on self-reports, which suffer from social desirability. We need more objective measures of information processing, such as eye movements, to effectively analyze the reading of news. To provide the research community the opportunity to study human behaviors in relation to news truthfulness, we propose the FakeNewsPerception dataset. FakeNewsPerception consists of eye movements during reading, perceived believability scores, questionnaires including Cognitive Reflection Test (CRT) and News-Find-Me (NFM) perception, and political orientation, collected from 25 participants with 60 news

* Corresponding author.
E-mail address: enkelejda.kasneci@uni-tuebingen.de (E. Kasneci).
Social media: (Ö. Sümer), (E. Bozkir), (S. Utz), (E. Kasneci)

https://doi.org/10.1016/j.dib.2021.106909

items. Initial analyses of the eye movements reveal that human perception differs when viewing true and fake news.

Specifications Table

Subject	Data Science
Specific subject area	Applied Machine Learning, Eye Movements, Fake News Perception,
Type of data	Table Image Figure
How data were acquired	Eye tracker computer (Tobii Pro-Spectrum), Tobii Pro-Lab software to record and export raw eye tracking data, and paper-based survey • Stimuli images are a part of the dataset. The questionnaire was conducted in German. English versions and German translations of tests, Cognitive Reflection Test (CRT) and News-Find-Me Perception (NFM), are included in the paper.
Data format	Analyzed Filtered
Parameters for data collection	Data is composed of participants' eye movements while viewing real and fake news stories, their ratings on the believability of stories, and post-tests including Cognitive Reflection Test (CRT), News-Find-Me (NFM), and perceived political orientation.
Description of data collection	The data collection procedure contains three parts. In the first part, participants read the news text without being informed about the actual purpose of the study regarding the true/fake news. Eye movement data collected from this part is reported in the dataset. In the second part of the experiment, self-reported believability scores are obtained from the participants by providing the same news texts and the Likert scales. In the last part of the experiment, participants complete the questionnaires of CRT, NFM, and self-reported political orientation along with demographic information.
Data source location	Department of Computer Science University of Tübingen
Data accessibility	Repository name: FakeNewsPerception: An Eye Movement Dataset on the Perceived Believability of News Stories Data identification number: https://doi.org/10.7910/DVN/C1UD2A Direct URL to data: https://doi.org/10.7910/DVN/C1UD2A

Value of the Data

- The FakeNewsPerception dataset contains eye movement data from participants reading news stories, perceived believability scores, and questionnaires, including the Cognitive Reflection Test (CRT) [4], News-Finds-Me perception (NFM) [2], and perceived political orientation. It can be used as a benchmark for scanpath classification and visualization.
- The majority of scanpath classification studies are focused on understanding the viewer's task as in Yarbus experiments.[1] However, the use of eye tracking for perceived believability is a

[1] Yarbus' experiment is accepted as the foundation of modern eye movement research. In his book, translated to English in 1967 [1], Yarbus reported that eye movements were task-dependent. His observer viewed a painting, I. E. Repin's *The Unexpected Visitor* (1884), for 21 minutes with optical stalks attached to the sclera. During visual search tasks such as "give the ages of the people," or "remember the clothes worn by the people", eye movements were fundamentally different from free-viewing.

new topic. Thus, FakeNewsPerception can be useful for computer scientists and cognitive psychologists developing scanpath representations and classifying either perceived believability or content type (true/false). It can also be useful for investigating the relationship between eye movements and various standardized questionnaire data.
- The dataset contains eye movement data and region of interest annotations to create a novel scanpath descriptor that represents discriminative features of eye movement behaviors.

1. Data Description

We provide our data in four data tables via four types of comma-separated values (CSV) files and describe as follows.

- Descriptive data file with questionnaire results (D1),
- Processed features file that includes aggregated features per participant and stimulus (D2),
- Processed data files per participant and news text, including timestamps, aggregated fixation information, saccades, and pupil diameter information (D3),
- Stimulus data including screenshots and annotated region of interests as visual content, headline, subheading, main text, and source of media (D4).

D1 is provided with a csv file. There are 13 columns and 26 rows. The first row represents feature types. Participant ID, gender, age, education, language competence, answers to cognitive reflection and news-find-me tests, and self-reported political orientation results are reported. If the questions were not answered, these fields were presented with NaN values. Table 1 depicts

Table 1
Questionnaire for Cognitive Reflection Test (CRT), News-Find-Me Perception (NFM), and political orientation. All questions were rated in 10-item Likert scale.

Item	Questions (in German)	Questions (in English)
crt1	Eine Suppe und ein Salat kosten insgesamt €5,50. Die Suppe kostet einen Euro mehr als der Salat, Wie viel kostet der Salat?	A soup and a salad cost a total of €5.50. The soup costs one euro more than the salad, How much is the salad?
crt2	Wenn 2 Krankenschwestern 2 Minuten brauchen, um den Blutdruck von 2 Patienten zu messen. Wie lange benötigen dann 200 Krankenschwestern, um den Blutdruck von 200 Patienten zu messen?	If 2 nurses take 2 min to measure the blood pressure of 2 patients. How long will it take 200 nurses to measure the blood pressure of 200 patients?
crt3	In einem See wachsen Seerosen. Jeden Tag verdoppelt sich die Menge der Seerosen. Die Seerosen brauchen 48 Tage, um den gesamten See zu bedecken. Wie lange würde es dauern, bis die Seerosen die Hälfte des Sees bedeckt haben?	Water lilies grow in a lake. Every day the amount of water lilies doubles. It takes 48 days for the water lilies to cover the entire lake. How long would it take for the water lilies to cover half of the lake?
nfm1	Ich verlasse mich darauf, dass meine Freunde mir sagen, was wichtig ist, wenn Nachrichten eintreten.	I rely on my friends to tell me what's important when news happens
nfm2	Ich kann gut informiert sein, auch wenn ich die Nachrichten nicht aktiv verfolge.	I can be well informed even when I don't actively follow the news
nfm3	Ich mache mir keine Sorgen darüber, dass ich auf dem Laufenden bleibe, weil ich weiß, dass die Nachrichten mich finden werden.	I don't worry about keeping up with the news because I know news will find me
nfm4	Ich verlasse mich auf Informationen von meinen Freunden, basierend auf Beiträgen, die sie auf sozialen Medien liken oder denen sie folgen.	I rely on information from my friends based on what they like or follow through social media
po	Wenn Sie an Ihre eigenen politischen Ansichten denken, wo würden Sie Ihre politische Grundhaltung einordnen?	When you think of your own political views, where would you classify your basic political stance?

the questions for cognitive reflection and news-find-me tests, and self-reported political orientation in English and German, together with column names in D1 per item. The first two items of CRT might be known by participants and, thus, were slightly adapted.

D2 is provided with a single csv file. This file reports participant and stimulus IDs, versions of the stimulus (real or fake), total view times, fixation and saccade counts, mean fixation and saccade durations, mean pupil diameters during stimulus viewing and during fixations, minimum and maximum durations for fixations and saccades, viewing time of Section 2, and the Likert scale measures indicating whether or not participants believed the truthfulness of the presented news. Eye tracking measurements were obtained from the first section of the experiment, whereas Likert scale measures were extracted from the second section. This file can be useful for statistical analyses between real and fake content or between different stimulus types.

D3 is provided for each participant and stimulus as a csv file. Each csv file includes aggregated data per stimulus for two eye movement types, namely fixations and saccades, which are important in the production of scanpath behaviors. There are 11 columns in the csv files and the first row represents features that are reported. These are the starting times of eye movement events, their durations, type of the event, unique event ID, mean pupil diameters during events, mean values of X and Y locations for fixations and starting, and ending X and Y positions of saccades. If a fixation event is presented in a line, starting and ending positions of saccades are reported with NaN. On the contrary, if the event type is saccade, mean values of X and Y locations are presented with NaN values.

In D4, screenshots of each stimulus along with the bounding box information for 5 different areas of stimulus with csv and json files are provided. The annotations are done using VIA annotation software [3]. Screenshots are provided with png files. The files (csv/json) that describe regions of stimuli consists of 7 different information categories including file name, file size, file attribute, number of annotated regions, region ID, bounding box attributes, and region type including image, headline, sub-headline, text, and source of news content. Transitions of the gaze between these areas can provide information on the believability of the stimuli. Fig. 2 depicts a sample screenshot of a news story and corresponding region of interest annotations, scanpath, and mouse movements.

Fig. 1 shows a representative image from the experimental setup used to collect eye-tracking data. Fig. 2 contains an example stimulus news text as well as ROI annotations (D4), scanpath data (D3), mouse movement, and click information.

Subsequent figures contain fundamental statistics of the dataset. The total viewing times, number of fixations, and number of saccades for fake and true content during Section 1 are depicted in Fig. 3. The distribution of rated believability of news in true and fake versions is in Fig. 4, and the distribution of perceived believability of each news story in true and fake versions is depicted in Fig. 5.

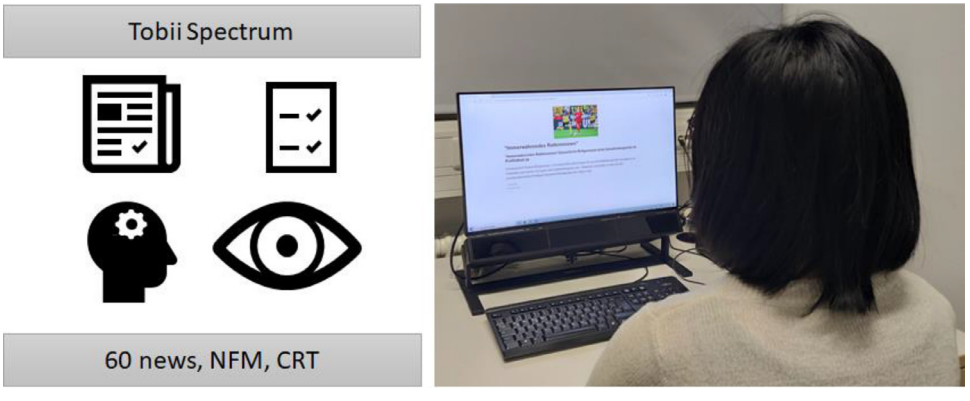
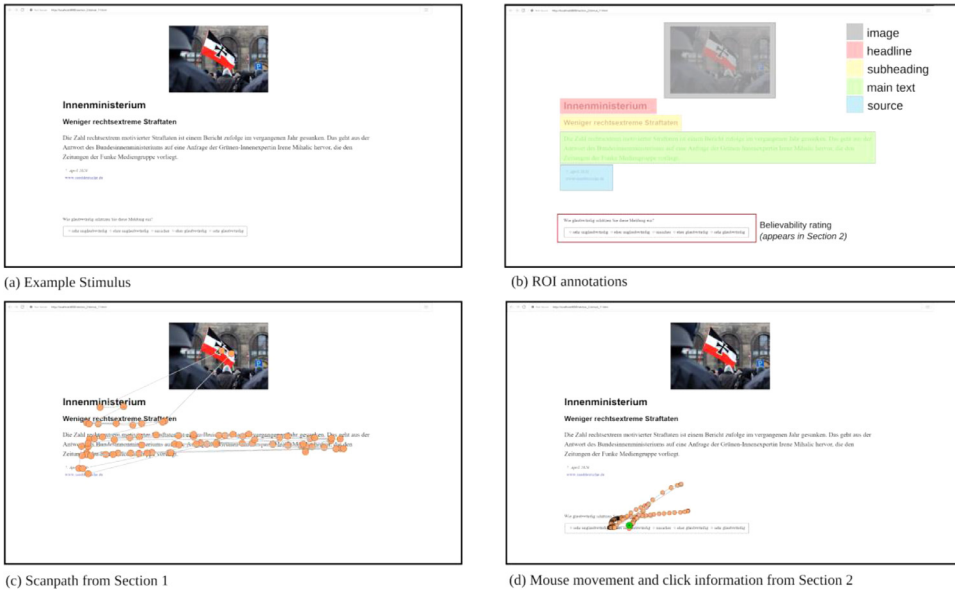


Fig. 1. An example from our data collection.



(a) Example Stimulus (b) ROI annotations

(c) Scanpath from Section 1 (d) Mouse movement and click information from Section 2

Fig. 2. Example stimulus new text, ROI annotations, scanpath data, mouse movement, and click information.

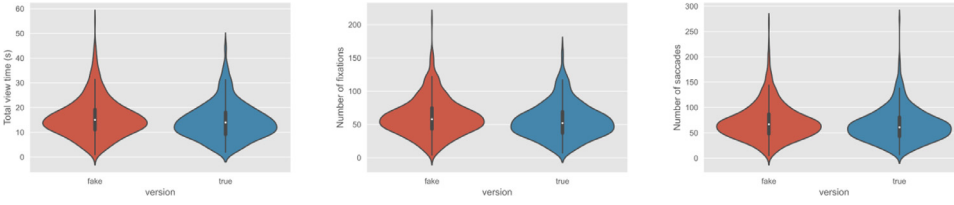


Fig. 3. Total viewing times, number of fixations, and number of saccades for fake and true content during section 1.

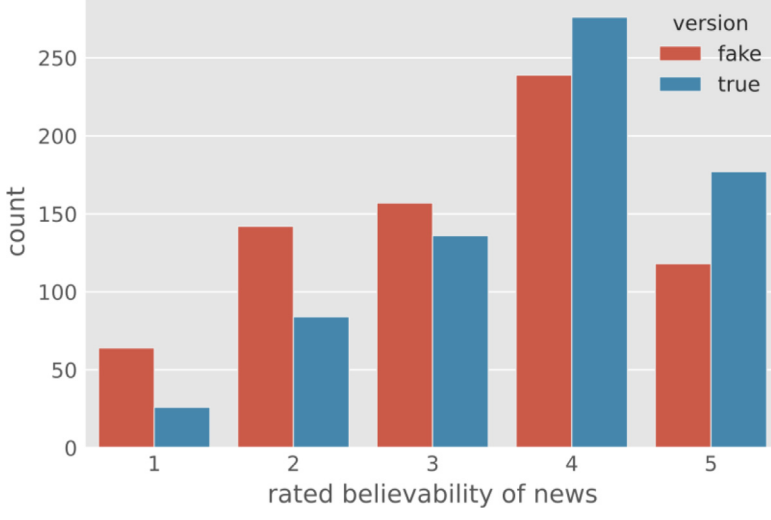


Fig. 4. The distribution of rated believability of news (5-item Likert scale, higher values represents more believable).

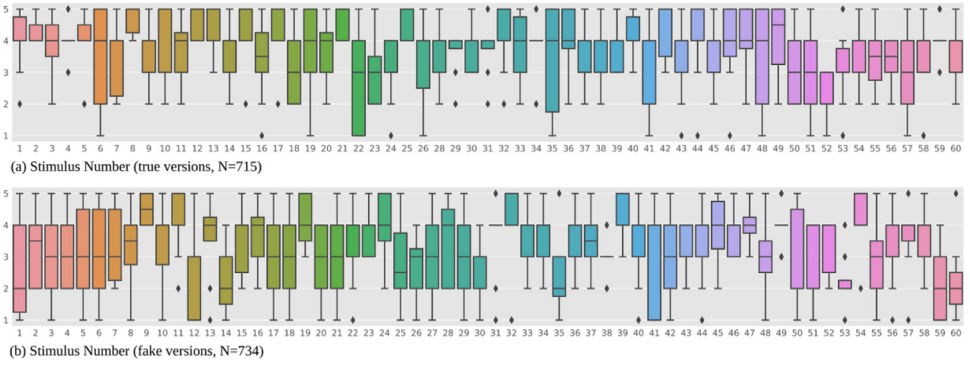


Fig. 5. The distribution of perceived believability of news. Except for small amounts of discarded data due to eye tracker issues, the number of samples per stimulus is balanced in true and fake groups (N-true=715, N-fake=734).

2. Experimental Design, Materials and Methods

2.1. Participants

A total of 27 participants were recruited for the experiment. All participants had normal or corrected-to-normal vision. Due to the sensory problems, we excluded two participants' entire data and used 25 participants (mean age=25.9, SD=4.8, 52% women, 48% men) in our further analysis. The participants were either students or researchers and recruited through the University's round email announcement. Their educational background was varying: high school graduates (#10), holding Bachelor's degree (#8), and Master's degree (#7). The ethics committee from the Leibniz-Institut für Wissensmedien in Tübingen approved our study procedures (approval \#2020–013). Every participant provided written informed consent for the collection and recording of their eye movement data and questionnaire information. Participants were informed and accepted the collected data, and results could be shared and published in an anonymous way. Participants were compensated \euro 10 for their participation in this study.

2.2. Apparatus

Eye movements were measured using a Tobii Spectrum desk-mounted eye tracker paired with a 23.8-inch screen (resolution of 1920 $\times$ 1080 pixels). The eye tracker sampled the gaze points at a rate of 600 Hz. We calibrated the eye tracker at the beginning of the experiment using a standard 9-point calibration; thus, there was no need for head stabilization. Furthermore, all curtains were closed, and the illumination of the room was controlled during the experiment. A representative image of how our experiment was done is given in Fig. 1.

2.3. Stimulus

The stimuli consisted of short text, separated in a header line and the news article. Next to the text, a related image was displayed (where possible from the original news article). We tried to resemble the typical look of a news article summary as presented on social media channels. In order to map the eye-tracking data more accurately to the content, the size of the stimulus and text was chosen to fill a large part of the screen.

2.4. Procedure and experimental design

To collect how the attentional mechanisms work when participants read news from a computer screen that can simulate reading from the print or website version of news content, we

designed a news reading study in the Department of Computer Science at the University of Tübingen. All participants were told the study was about reading comprehension. They were not, however, explicitly informed that there might be false content among the stimuli they viewed.

Our experiment procedure is described in the following. First, participants read the instructions and provided written informed consent. After successful eye tracking calibration (lower than a mean angular error of 0.50°), participants started the actual experiment which took about 40 min. Participants were free to pause or stop the experiment at any point. In the latter case, the collected data was deleted and not included in further analysis.

Our experiment contained three sections. In the first section, participants were asked to read 60 news items at their usual reading pace. Two versions of each text were provided, one containing true content from differing media sources and the other containing falsified information. In order to eliminate the effect of ordering, topics were presented in random order. We used the reversed version of each news text (true/falsified) for every other participant, thus presenting both versions of each stimulus the same number of times.

In the second section, participants read the same content a second time, but the viewing time was limited to 10 s. After 10 s, they were asked to rate the credibility of the news content on a 5-item Likert scale ("How credible do you rate this news?"). At the end of the two sections on the eye tracker computer, participants completed the CRT measure [4], "News-Finds-Me perception" [2], and a question regarding political orientation on a 10-item Likert scale on paper.

2.5. Data pre-processing

We used the Tobii Pro-Studio software to design and record our experiment and to pre-process the recorded eye tracking data. Eye movement events, i.e., fixations and saccades, were detected with a Tobii I-VT filter [5] with a minimum fixation duration of 60 ms and a velocity threshold of 30°/s. We included the samples where the participant's left and right eyes had valid data points more than 80% of a stimulus's viewing time. This threshold was defined to eliminate news story-participant pairs where the total success ratio of the eye-tracker was very low and we could not extract eye movement features reliably. 89.5% of all news stories were above this threshold and the mean and standard deviation of the success ratio used were $0.95 +/- 0.04$.

Fixations correspond to periods when eyes are stationary within an area of the stimuli and usually last between 200 ms and 350 ms [6]. The information extracted from the fixations corresponds to different aspects of the stimuli and content. For instance, longer fixations correspond to difficulties in extracting content or, in contrast, content that is more engaging [7]. Thus, reported fixations can be filtered further using higher fixation thresholds to analyze different aspects.

In contrast to fixations, saccades are high speed eye movements that correspond to transitions between fixations and, due to their involuntarily nature, are linked to various cognitive processes [8,9]. While having a greater number of saccades may indicate more searching, saccades that have significant deviations in terms of angles may illustrate a change in a user's goal [10]. In addition, saccade trajectories can be affected by the spatial locations of distractors [11].

In addition to eye movement types and parameters, we included pupillometry data in our dataset. We obtained pupil diameter values from the eye tracker directly. Additionally, we applied the Savitzky–Golay filter [12] (2nd order polynomial, a time window of 15 data points) to smooth the raw data values. As the stimuli content is randomized, we used the period of time when the participants viewed the initial instruction screen (the median value of only the first 10 data points corresponding to 15 ms) as divisive baseline correction [13]. In our analysis, we used the first section (free-viewing) and split the eye tracking data by the displayed news text.

Pupil diameter has been used to assess cognitive load [14]. Task difficulty [15], emotional arousal [16], fatigue [17], and pain [18] have also been studied using pupil diameter measurements. However, pupil dilations can occur due to different illumination conditions.

2.6. Technical validation

To validate our experimental setup, we ran initial analysis on total viewing time, number of fixations, and number of saccades between real and fake content using the data obtained in the first section of our experiment. Total viewing time corresponds to the duration between starting timestamp and ending timestamp per stimulus. We applied the Shapiro-Wilk and Levene tests to understand whether distributions were normally distributed and that the homogeneity of variance was preserved, respectively. Since we obtained non-normal distributions, we applied the Wilcoxon rank sum test using α level of 0.05. We found that participants attended significantly more time on the fake content ($M = 15.9$ s, SD=7.8 s) than on the true content ($M = 14.6$ s, SD=7.4 s) with $p = 5.4E\text{-}4$ as depicted in Fig. 3(a). In addition, the number of fixations and saccades differed significantly. Participants had a significantly higher number of fixations on the fake content ($M = 61.01$, SD=27.89) than on the true content ($M = 55.63$, SD=26.48) with $p = 9.0E\text{-}5$. Furthermore, participants had a higher number of saccades on the fake content ($M = 71.24$, SD=34.46) than on the true content ($M = 64.89$, SD=32.11) with $p = 2.56E\text{-}4$. The numbers of fixations and saccades for fake and true content are depicted in Figs. 3(b), and 3(c), respectively. Even though the participants were not informed as to whether they were reading fake or true content, their attention times, number of fixations, and saccades differed. This indicates that further fine-grained analysis can offer additional insights on how trust or doubt in the news item is constructed.

Ethics Statement

The ethics committee from the Leibniz-Institut für Wissensmedien in Tübingen approved our study procedures (approval \#2020–013) and the informed consent was obtained for experimentation with human subjects.

CRediT Author Statement

Ömer Sümer, Efe Bozkir, Thomas Kübler, Sven Grüner, Sonja Utz, and **Enkelejda Kasneci:** conceived the experiment; **Sven Grüner:** prepared the initial news items; **Ömer Sümer:** conducted the experiment; **Ömer Sümer** and **Efe Bozkir:** processed the data and analyzed the results. All authors prepared and reviewed the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have or could be perceived to have influenced the work reported in this article.

Acknowledgments

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC–Number 2064/1, Project number 390727645. We acknowledge support by Open Access Publishing Fund of University of Tübingen.

References

[1] A. Yarbus, Eye Movements and Vision, B. Haigh, Trans., Plenum Press New York, 1967.
[2] H. Zúñiga, B. Weeks, A. Ardèvol-Abreu, Effects of the news-finds-me perception in communication: social media use implications for news seeking and learning about politics, J. Comput.-Mediat. Commun. 22 (3) (2017) 105–123.
[3] A. Dutta, A. Zisserman, The VIA annotation software for images. audio and video, in: Proceedings of the 27th ACM International Conference on Multimedia, Association for Computing Machinery, 2019, pp. 2276–2279.

[4] S. Frederick, Cognitive reflection and decision making, J. Econ. Perspect. 19 (4) (2005) 25–42.
[5] Olsen, A., The Tobii I-VT fixation filterTobii Technology, (2012) 1–21.
[6] K. Rayner, Eye movements in reading and information processing: 20 years of research, Psychol. Bull. 124 (3) (1998) 372.
[7] M. Just, P. Carpenter, Eye fixations and cognitive processes, Cogn. Psychol. 8 (4) (1976) 441–480.
[8] S. Liversedge, J. Findlay, Saccadic eye movements and cognition, Trends Cogn. Sci. (Regul. Ed.) 4 (1) (2000) 6–14.
[9] S. Hutton, Cognitive control of saccadic eye movements, Brain Cogn. 68 (3) (2000) 327–340.
[10] J. Goldberg, X. Kotval, Computer interface evaluation using eye movements: methods and constructs, Int. J. Ind. Ergon. 24 (6) (1999) 631–645.
[11] S. Stigchel, J. Theeuwes, Relation between saccade trajectories and spatial distractor locations, Cogn. Brain Res. 25 (2) (2005) 579–582.
[12] A. Savitzky, M. Golay, Smoothing and differentiation of data by simplified least squares procedures, Anal. Chem. 36 (1964) 1627–1639.
[13] S. Mathot, Safe and sensible preprocessing and baseline correction of pupil-size data, Behav. Res. Methods 50 (1) (1964) 94–106.
[14] T. Appel, N. Sevcenko, F. Wortha, K. Tsarava, K. Moeller, M. Ninaus, E. Kasneci, P. Gerjets, Predicting cognitive load in an emergency simulation based on behavioral and physiological measures, in: 2019 International Conference on Multimodal Interaction, 2019, pp. 154–163.
[15] T. Appel, C. Scharinger, P. Gerjets, E. Kasneci, Cross-subject workload classification using pupil-related measures, In Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications (2018) 1–8, doi:10.1145/3204493.3204531.
[16] R. Henderson, M. Bradley, P. Lang, Emotional imagery and pupil diameter, Psychophysiology 55 (6) (2018) e13050.
[17] J. Hopstaken, D. Linden, A. Bakker, M. Kompier, The window of my eyes: task disengagement and mental fatigue covary with pupil dynamics, Biol. Psychol. 110 (2015) 100–106.
[18] D. Wildemeersch, N. Peeters, V. Saldien, M. Vercauteren, G. Hans, Pain assessment by pupil dilation reflex in response to noxious stimulation in anaesthetized adults, Acta Anaesthesiol. Scand. 62 (8) (2018) 1050–1056.