

Can Eye Movement Synchronicity Predict Test Performance With Unreliably-Sampled Data in an Online Learning Context?

Marian Sauter*
sauter.marian@gmail.com
Ulm University, Institute of
Psychology and Education
Ulm, Germany

Susanne Hummel
susanne-1.hummel@uni-ulm.de
Ulm University, Institute of
Psychology and Education
Ulm, Germany

Teresa Hirzle*
teresa.hirzle@uni-ulm.de
Ulm University, Institute of Media
Informatics
Ulm, Germany

Enrico Rukzio
enrico.rukzio@uni-ulm.de
Ulm University, Institute of Media
Informatics
Ulm, Germany

Tobias Wagner
tobias.wagner@uni-ulm.de
Ulm University, Institute of Media
Informatics
Ulm, Germany

Anke Huckauf
anke.huckauf@uni-ulm.de
Ulm University, Institute of
Psychology and Education
Ulm, Germany

ABSTRACT

Webcam-based eye-tracking promises easy and quick data collection without the need for specific or additional eye-tracking hardware. This makes it especially attractive for educational research, in particular for modern formats, such as MOOCs. However, in order to fulfill its promises, webcam-based eye tracking has to overcome several challenges, most importantly, varying spatial and temporal resolutions. Another challenge that the educational domain faces especially, is that typically individual students are of interest in contrast to average values. In this paper, we explore whether an attention measure that is based on eye movement synchronicity of a group of students can be applied with unreliably-sampled data. Doing so we aim to reproduce earlier work that showed that, on average, eye movement synchronicity can predict performance in a comprehension quiz. We were not able to reproduce the findings with unreliably-sampled data, which highlights the challenges that lie ahead of webcam-based eye tracking in practice.

CCS CONCEPTS

• **Applied computing** → **E-learning**.

KEYWORDS

online teaching, webcam, eye-tracking, remote learning

ACM Reference Format:

Marian Sauter, Teresa Hirzle, Tobias Wagner, Susanne Hummel, Enrico Rukzio, and Anke Huckauf. 2022. Can Eye Movement Synchronicity Predict Test Performance With Unreliably-Sampled Data in an Online Learning Context?. In *2022 Symposium on Eye Tracking Research and Applications (ETRA '22)*, June 8–11, 2022, Seattle, WA, USA. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3517031.3529239>

*Both authors contributed equally to the paper



This work is licensed under a Creative Commons Attribution International 4.0 License.

ETRA '22, June 8–11, 2022, Seattle, WA, USA
© 2022 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9252-5/22/06.
<https://doi.org/10.1145/3517031.3529239>

1 INTRODUCTION

Webcam-based eye-tracking can exploit the advantages of online recruitment methods [Sauter et al. 2020], enable researchers in conducting more resource-efficient studies, and recruit from a potentially more diverse participant pool. In educational research, there has been a recent pioneering webcam-based eye-tracking study by Madsen et al. [2021]. They show that the synchronicity between eye movements can predict a subsequent test performance. The authors suggest that if any given student shows similar eye movement patterns to the 'average' student, this reveals their attentiveness. They further claim that "*Our results suggest that online education could be made adaptive to a student's level of attention in real time.*" This is a bold, yet promising, claim given that this entails their predictive algorithm needs to work not only for students on average but for any individual student. Additionally, real-time functionality requires the algorithm to be able to deal with less-than-ideal data, as no long averages can be calculated.

In the present paper, we explore whether this claim holds for non-ideal webcam-based eye tracking data. We present results of a webcam-based eye tracking study and analyse whether the synchronicity-based attention measure can be obtained from non-ideal data, i.e., data with a low temporal resolution. To implement adaptive learning material that can be presented to students, the attention measure has to be validated in a real-life scenario. This includes students that have standard laptops that might not be able to provide high temporal resolution of eye tracking measures. In summary, we aimed to reproduce Madsen et al.'s [2021] finding on the synchronicity of eye movements and test performance in a non-ideal eye tracking environment. To that end, we conducted an eye tracking study with the Labvanced [Scicoverly GmbH 2022] webcam-based eye tracking algorithm. We presented three videos to a set of 27 study participants and asked them comprehension questions about the videos. We then evaluated whether their eye movement synchronicity predicted test performance. Our results reveal that it is difficult to obtain correlations between eye movement synchronicity and test performance in a webcam-based eye tracking study without specific eye-tracking equipment. This suggests that building adaptive learning algorithms that are based on

individual students' eye movements is still challenging in settings using standard hardware only.

2 RELATED WORK

Webcam-based eye-tracking can be used to pilot eye-tracking data, which could be beneficial especially when labs are closed such as during the Covid-19 pandemic [Sauter et al. 2020; Yang and Krajbich 2020]. In order to record webcam gaze data, there are not many software solutions available. The aforementioned and other earlier studies used a Javascript-based open source script WebGazer [Papoutsaki et al. 2016]. The algorithm estimates the gaze position by evaluating the user's video feed. Additionally, the model is simply trained by mouse clicks, which rests on evidence indicating that, in general, we look at the cursor when we click somewhere [Huang et al. 2012]. Notably, Webgazer is still in active development [Huang 2022]. In addition to WebGazer, proprietary online solutions have emerged. Of most interest for behavioral scientists are arguably the built-in solutions from existing and already-used experimental platforms such as Gorilla [Science 2022] and LabVanced [Scicoverly GmbH 2022]. While Gorilla's algorithm is based on WebGazer, LabVanced [Scicoverly GmbH 2022] offers an own proprietary solution with limited information on how the algorithm works.

So far, only selected studies have employed remote webcam-based eye tracking [Federico and Brandimonte 2019; Lin et al. 2022; Madsen et al. 2021; Robal et al. 2018a; Schneegans et al. 2021; Schröter et al. 2021; Semmelmann and Weigelt 2018; Zhao et al. 2017]. One of the first studies directly contrasting webcam-based and conventional high resolution infrared-based data was conducted by Semmelmann and colleagues [Semmelmann and Weigelt 2018] who compared the precision of the data in three tasks (fixation, pursuit, and free viewing) and found that estimates are slightly less precise with a higher variance in webcam-based eye tracking but the method was found to be suitable for all three tasks. Importantly, they note that a disadvantage of webcam-based eye tracking is the potentially high calibration effort which can take a considerable amount of time, for their experiments even up to 50% of the total study time. Additionally, participants might drop out due to technical requirements such as processing power [Yang and Krajbich 2020] or webcam resolution [Scicoverly GmbH 2022].

In the last years, more and more educational videos incorporated the video feed of a person (an 'on-screen instructor') [Henderson and Schroeder 2021]. It is typically shown in the corner of the screen, most often in the upper right. A recent systematic review summarized the impact on-screen instructors have on the learning outcomes [Henderson and Schroeder 2021], with mixed results ([Henderson and Schroeder 2021], Table 1). Some students reported that the on-screen instructor could have impaired their ability to focus properly [Henderson and Schroeder 2021; Yu 2021]. Before, it was already shown that videos of humans, attract eye movements and capture attention [Bindemann et al. 2005; Langton et al. 2008; Wang et al. 2020]. It stands to reason that, on-screen instructors may influence the synchronicity between students, as some are more distracted (i.e. larger variance in eye movements) with an instructor present. In the present study, we used the LabVanced webcam-based eye-tracking algorithm as it is marketed as superior to WebGazer.

However, independent estimates about this proprietary algorithm's precision and practicability are hard to obtain. In particular, we tried reproducing Madsen et al.'s [2021] finding on the synchronicity of eye movements and test performance. Additionally, we explored potential presenter-related effects. Similarly, we set out to get hold of the data quality using this particular eye tracking algorithm.

3 METHODS

3.1 Online Eye Tracking Study with Students

We conducted an online eye tracking study to gather the eye movement data based on educational video clips.

3.1.1 Study Design. The study was conducted as an online study using the survey platform Prolific.co for recruiting participants. We only recruited students as participants. For the implementation of the study we used Labvanced [Scicoverly GmbH 2022]— a cloud-based solution for conducting online experiments, which includes a web cam-based eye tracking module. This study aimed to gather realistic eye movement data of students who watched the online videos. In the study, each participant watched three out of six videos in a randomized order. We chose to let participants watch only three of the videos to keep the study duration relatively short. If they had to watch all six videos, we expected a drop in attention.

3.2 Online Lecture Videos

We evaluated the gaze visualizations based on six videos. We chose the videos from a set of explanatory films lasting less than 5 minutes each. Four of the six videos we chose, were recorded for an online psychology conference. To assure comprehensibility of the conference videos, two undergraduate authors of this paper independently rated the videos' comprehensibility on a 4-point scale (1: "very well comprehensible" to 4: "not at all comprehensible"). From all videos with author consent which were rated as "very well comprehensible" by both raters, we selected four videos; two conference videos with a female presenter, two conference videos with a male one; two of those with a visible presenter, and two in which a presenter was not visible. The other two videos were recorded e-learning videos (on enzymes¹ and planet measurements²), that have already been used in [Madsen et al. 2021]. We decided to use these two videos out of many more, because they were most comparable to the conference videos in terms of their length and structure. All presentations were held in English.

3.2.1 Procedure. At first, the participants read the informed consent of participation. They had to agree to the informed consent and to the recording and publishing of their data. If they did not agree, they were automatically redirected to the end of the study. If they did agree, the participants started the study by indicating their experience with online learning. They were asked if they were currently enrolled at an university and whether they had attended any online lectures in the last semester, how many lectures they attended were held online, and for how long the online lectures usually lasted. After that, they were introduced into the procedure of the following study and performed a short sound check to adjust their audio volume. They then performed a calibration procedure


¹<https://www.youtube.com/watch?v=lkRZKqDdwzU>, last accessed: February 17, 2022

²<https://www.youtube.com/watch?v=bYgV9nvgJ3E>, last accessed: February 17, 2022

Methods

- 2 experimental conditions: security vs. universalism framing

„Better husbandry conditions... [...] no more meat scandals“



„Better husbandry conditions... [...] no more cruelty to animals“

- 1 continuous moderator: egoistic, altruistic and biospheric value scales (de Groot & Steg, 2007; 2008)
 - Egoistic (5 Items), $\alpha = .76$
 - Altruistic (4 Items), $\alpha = .74$
 - Biospheric (4 Items), $\alpha = .88$

Eschert & Nuszbaum | TealP 2021

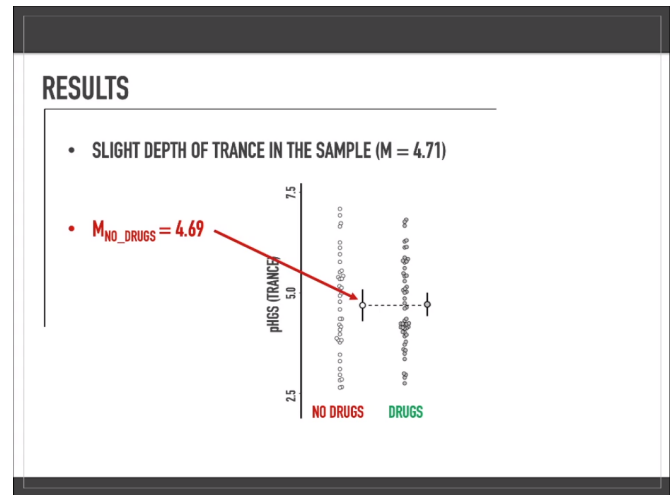


Figure 1: Screenshots of two of the presentation videos that were used in the study. The presenters of these videos gave their consent that we could use a screenshot of their videos in our paper.

provided by Labvanced, which lasted eight minutes. In the calibration procedure, the participants were instructed to position their head in a specific orientation while fixating between eight and 15 scattered points on the screen successively. Each point had to be fixated for three seconds. Overall, participants conducted ten calibration trials. In the first two trials, participants were required to center their heads and to look forward. Then, participants conducted four different head shifts, and four different head tilts in a calibration trial, respectively. This allows eye tracking even if the participants perform subtle head movements during the study. After completing the calibration procedure, the participants watched the three presentation videos. The first two videos were two out of the four conference videos, which were presented counterbalanced. The third video was randomized one of the two recorded e-learning videos. For each video, they had to answer a knowledge test consisting of six questions and an attention check. The attention check was used to determine whether participants paid attention to the videos and the questionnaires (i.e., “Click on the option green in the list below.”). We used the knowledge test to ensure that the participants watched and understood the content of the videos. Additionally, the participants had to rate three more items after each video. The two items “I was focused on watching the videos most of the time.” and “For me, the learning content was easy to understand.” had to be rated on a 5-point scale reaching from “totally agree” to “totally disagree”, indicating the focus and understanding of the participants. The third item “How similar was the video to online lectures you recently attended?” was rated on a 5-point scale reaching from “Not at All Similar” to “Very Similar”, indicating the similarity of the videos to the participants’ online lectures.

3.2.2 Eye Tracking. While the participants watched the videos, their gaze position (x, y) on the screen was recorded together with a confidence value (c , range: $0 - 1$), indicating the quality of the tracking. We used the built-in webcam-based eye tracking method of Labvanced that uses deep learning to estimate the participants’

gaze points on the screen. The sampling rate of the recorded data varied among participants from 4 Hz to 13 Hz (mean: 7 Hz, SD: 20 Hz), depending on the technical properties of their device. The gathered raw eye tracking data was critically appraised by means of the gaze confidence value provided by Labvanced for each gaze coordinate (a value between 0 and 1, with 1 indicating high confidence). The confidence value is calculated by an algorithm which takes head movements into account: the more the head is moving, the smaller, the confidence estimation.

4 RESULTS

4.1 Participants and Demographics

Overall, 89 participants initialized the study via the platform Prolific. Of those, 32 observers produced at least partial data in the system (i.e. they managed to do the initial eye tracking calibration) and 27 of those have complete datasets. Of those, participants were on average 22 years old (range: 19-38 years; 21 female, 6 male) and indicated a diverse range of 15 nationalities, most notably Portugal (5), United Kingdom (4) and Poland (3). Those who completed took 23 minutes on average (median). Two were further excluded for not having passed the attention check.

4.2 Average Confidence

For all participants, the average confidence in the gaze samples was 72% (57%-96%).

4.3 Further Preparation of Gaze Data

The study in [Madsen et al. 2021] uses webcam-based eye tracking similarly with educational videos. They claim to have a consistent sampling rate of 30Hz using WebGazer. In their study, eye movement synchronization is calculated as follows: For each time point the fixation distance of each observer to the median fixation from all observers is computed in the x-axis and y-axis separately. They then calculate an additional velocity vector, i.e. how fast was the

Table 1: Regression results with the predictors inter-subject correlation (ISC; Model 1) and additionally presenter condition (Model 2) and results with the predictors synchronicity and relative gaze samples on the presenter (Model 3) and the outcome test performance (in %-correct).

	<i>Dependent variable:</i>		
	Model 1 (1)	Model 2 (2)	Model 3 (3)
ISC	-0.004(0.002)	0.001 (0.002)	-0.013* (0.006)
presenter [corner]		0.165* (0.088)	
presenter [full]		-0.133 (0.112)	
ISC:presenter [corner]		-0.006 (0.004)	
ISC:presenter [full]		-0.006 (0.004)	
gazeOnPresenter			-0.022 (0.028)
ISC:gazeOnPresenter			0.002 (0.001)
Constant	0.707*** (0.053)	0.689*** (0.054)	0.925*** (0.122)
Observations	25	77	16
R ²	0.096	0.428	0.360
Adjusted R ²	0.057	0.388	0.200
Residual Std. Error	0.128 (df=23)	0.190 (df=71)	0.169 (df=12)
F Statistic	2.451 (df=1; 23)	10.622*** (df=5; 71)	2.253 (df=3; 12)

Note:

*p<0.1; **p<0.05; ***p<0.01

eye movement leading to this fixation and compare this to the median velocity vector across observers as well. They then combine the three values (x, y and v) to one variable in a weighted fashion (weighted inter-subject correlation): $wISC = w_1 * x + w_2 * y + w_3 * v$. The weights are chosen accordingly to maximize prediction performance, see [Madsen et al. 2021] for details.

The authors also looked at which sampling frequency lets the wISC correlate best to test performance and concluded that 1 Hz is optimal (see their figure). For our data, we incorporated this finding: First, for each video, we defined 1-second windows. Second, for each participant and video, we aggregated their samples for each window (mean x/y values). We also calculated the mean fixation position across all observers for each video and time window. Note that not all observers have data for all windows.

As we cannot calculate gaze velocity without having recorded all saccades (potentially), we only calculated its distance component. This means that we calculated the point distance between each two bins and note this as the 'minimally traveled distance'. Accordingly, all three measurements were aggregated with equal weights: $ISC = x + y + d$.

4.4 Linear Regressions

We then computed linear regressions (see Table 1) using the predictor variable regressing on mean correctness (Model 1). As noted above, there might be presenter effects on the gaze distributions, so we included this factor as well (Model 2).

4.5 Area of Interest Analysis

Looking at the distribution of the gaze samples across the slide, we calculated the percentage of samples falling within the presenter area (i.e. the area of the whole video feed, not only the presenters face) for the two videos where the presenter was in the corner. For

one video, no gaze samples lied within the specified area (the presenter area was quite small). For the other video, across participants on average 4.92% (range 0% - 27%) of samples lied within the presenter area. We calculated a linear regression (see 1, Model 3) with the percentage of samples in the presenter area and the interaction with the synchronicity measure, but found no significant effects.

5 DISCUSSION

In the present study, we evaluated the potential of using unreliably-sampled eye tracking data through webcams to correlate eye movement synchronicity with test performance. In analysing our data, we chose to adopt simpler methods as compared to [Madsen et al. 2021]. We specifically explored whether a simple measure of synchronicity can inform subsequent test performance. We could not show this beyond doubt (only for the two videos in which the presenter was in the corner, there seems to be a slight trend). The generated synchronicity measure does not seem to be of predictive value across all video types, as was the case in [Madsen et al. 2021]. Reasons for this may lie in Madsen et al.'s [2021] specific approach to calculate the ISC or pre-process data (e.g. remove all participants with less than 15 Hz sampling rate), that we could not reproduce with the available data. As an side finding, we observed that there might be indications, that a presenter being visible in the corner of a presentation slide might increase the test performance. However, this can only be considered preliminary and needs to be experimentally manipulated, as currently, we did not compare the same videos in the presenter on/off conditions.

We found that the estimated confidence in the gaze samples was around 72%. When the data had been cleaned for unreliable sample, we would have lost nearly a fourth of trials. However, to do this in an informed manner, it is crucial to understand how this confidence

is estimated, which LabVanced does not fully disclose. It is just mentioned that the measure includes a head movement correction.

While there are some examples of successfully demonstrated effects using webcam-based eye tracking, it may still be poor in spatial and especially in temporal resolution (typically max. 30Hz compared to 1000Hz in the laboratory). Of course, this also affects the current data of our study in which we gathered the eye movement data of the students. Thus, a common issue with online eye tracking studies [Robal et al. 2018b], the data is characterized with a low and varying eye-tracking sampling rate for different participants.

5.1 Opinion: Personal Experience with the LabVanced Eye Tracking

This subsection is meant as a starting point of discussion and therefore clearly labeled as opinion. In our study, the LabVanced eye tracking algorithm produced unreliable results. First of all, many observers were not even able to finish the calibration (see 4.1). We received individual reports of participants being 'stuck' in the calibration procedure with no way out. But even calibration procedure did not guarantee complete datasets, as participants dropped out mid-Experiment, seemingly related to their processing power. Notably, in our data collection period (for this study and two other student projects) from April to June 2021, the interface and the LabVanced eye tracking algorithm was re-worked twice without a warning. We do not know, whether the currently-implemented version of the eye-tracker works more reliably. A reason for this is also because the 'white paper' describing the algorithm has not been released yet. We suggest general caution with using proprietary online tools with no demonstrated backwards-compatibility.

6 CONCLUSION

Webcam-based eye tracking holds many promises for eye-tracking studies in the educational domain as the external validity could be dramatically increased when standard webcams can be used reliably. It also allows researchers to pilot approaches "in-the-wild" and outside the lab. However, to fulfill these promises reliably, challenges such as unreliably-sampled eye movement data have to be overcome for all major solutions. One concrete challenge is whether webcam-based eye tracking can be applied on an individual basis, i.e., does it work beyond average values? We presented a study that tested whether an attention measure based on eye movement synchronicity can predict test performance based on average values. Our results did not suggest a correlation between eye movement synchronicity and test performance. However, we found that considering various kinds of presentation formats, especially concerning the depiction of a teacher, seems to be of importance.

ACKNOWLEDGMENTS

This project is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, 425867974, to Anke Huckauf and Enrico Rukzio) and is part of Priority Program SPP2199 Scalable Interaction Paradigms for Pervasive Computing Environments.

REFERENCES

Markus Bindemann, A. Mike Burton, Ignace T. C. Hooge, Rob Jenkins, and Edward H. F. de Haan. 2005. Faces retain attention. *Psychonomic Bulletin & Review* 12, 6

- (2005), 1048–1053. <https://doi.org/10.3758/bf03206442>
- Giovanni Federico and Maria A Brandimonte. 2019. Tool and object affordances: An ecological eye-tracking study. *Brain and cognition* 135 (2019), 103582.
- Marsha L. Henderson and Noah L. Schroeder. 2021. A Systematic review of instructor presence in instructional videos: Effects on learning and affect. *Computers & Education Open* 2 (2021), 100059. <https://doi.org/10.1016/j.caeo.2021.100059>
- Jeff Huang. 2022. *WebGazer.js*. WebGazerGit. <https://github.com/brownhci/WebGazer>
- Jeff Huang, Ryen White, and Georg Buscher. 2012. User See, User Point: Gaze and Cursor Alignment in Web Search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) (CHI '12). Association for Computing Machinery, New York, NY, USA, 1341–1350. <https://doi.org/10.1145/2207676.2208591>
- Stephen R. H. Langton, Anna S. Law, A. Mike Burton, and Stefan R. Schweinberger. 2008. Attention capture by faces. *Cognition* 107, 1 (2008), 330–342. <https://doi.org/10.1016/j.cognition.2007.07.012>
- Zhongjun Lin, Yinghua Liu, Hui Wang, Zhengji Liu, Sijin Cai, Zhicheng Zheng, Yongjin Zhou, and Xinyu Zhang. 2022. An eye tracker based on webcam and its preliminary application evaluation in Chinese reading tests. *Biomedical Signal Processing and Control* 74 (2022), 103521. <https://doi.org/10.1016/j.bspc.2022.103521>
- Jens Madsen, Sara U. Júlio, Paweł J. Gućik, Richard Steinberg, and Lucas C. Parra. 2021. Synchronized eye movements predict test scores in online video education. *Proceedings of the National Academy of Sciences of the United States of America* 118, 5 (2021), 9 pages. <https://doi.org/10.1073/pnas.2016980118>
- Alexandra Papoutsaki, Patsorn Sangkloy, James Laskey, Nediya Daskalova, Jeff Huang, and James Hays. 2016. WebGazer: Scalable Webcam Eye Tracking Using User Interactions. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI-16)*. AAAI, AAAI, -, 3839–3845.
- Tarmo Robal, Yue Zhao, Christoph Lofi, and Claudia Hauff. 2018a. Webcam-Based Attention Tracking in Online Learning: A Feasibility Study. In *23rd International Conference on Intelligent User Interfaces* (Tokyo, Japan) (IUI '18). Association for Computing Machinery, New York, NY, USA, 189–197. <https://doi.org/10.1145/3172944.3172987>
- Tarmo Robal, Yue Zhao, Christoph Lofi, and Claudia Hauff. 2018b. Webcam-Based Attention Tracking in Online Learning: A Feasibility Study. In *23rd International Conference on Intelligent User Interfaces* (Tokyo, Japan) (IUI '18). Association for Computing Machinery, New York, NY, USA, 189–197. <https://doi.org/10.1145/3172944.3172987>
- Marian Sauter, Dejan Draschkow, and Wolfgang Mack. 2020. Building, Hosting and Recruiting: A Brief Introduction to Running Behavioral Experiments Online. *Brain Sciences* 10, 4 (2020), 251. <https://doi.org/10.3390/brainsci10040251>
- Tim Schneegans, Matthew D. Bachman, Scott A. Huettel, and Hauke Heekeren. 2021. Exploring the Potential of Online Webcam-based Eye Tracking in Decision-making Research and Influence Factors on Data Quality. *PsyArXiv* -, - (2021), 81 pages. <https://doi.org/10.31234/osf.io/zm3us>
- Iris Schröter, Nico Rolf Grillo, Margarethe Kristine Limpak, Bilel Mestiri, Benedikt Osthold, Fourat Sebt, and Marcus Mergenthaler. 2021. Webcam Eye Tracking for Monitoring Visual Attention in Hypothetical Online Shopping Tasks. *Applied Sciences* 11, 19 (2021), 9281. <https://doi.org/10.3390/app11199281>
- Scicoverly GmbH. 2022. *LabVanced*. LabVanced. <http://labvanced.com>
- Cauldron Science. 2022. *Gorilla*. Gorilla. <http://gorilla.sc>
- Kilian Semmlmann and Sarah Weigelt. 2018. Online webcam-based eye tracking in cognitive science: A first look. *Behavior Research Methods* 50, 2 (2018), 451–465. <https://doi.org/10.3758/s13428-017-0913-7>
- Jiahui Wang, Pavlo Antonenko, and Kara Dawson. 2020. Does visual attention to the instructor in online video affect learning and learner perceptions? An eye-tracking analysis. *Computers & Education* 146 (2020), 103779. <https://doi.org/10.1016/j.compedu.2019.103779>
- Xiaozhi Yang and Ian Krajbich. 2020. Webcam-based online eye-tracking for behavioral research. *PsyArXiv* -, - (2020), 31 pages. <https://doi.org/10.31234/osf.io/qhme6>
- Zhonggen Yu. 2021. The effect of teacher presence in videos on intrinsic cognitive loads and academic achievements. *Innovations in Education and Teaching International* 0, 0 (2021), 1–12. <https://doi.org/10.1080/14703297.2021.1889394> arXiv:<https://doi.org/10.1080/14703297.2021.1889394>
- Yue Zhao, Christoph Lofi, and Claudia Hauff. 2017. Scalable Mind-Wandering Detection for MOOCs: A Webcam-Based Approach. In *Data Driven Approaches in Digital Education*. Springer International Publishing, Cham, 330–344.