A gaze-based study design to explore how competency evolves during a photo manipulation task

Nora Castner
nora.castner@uni-tuebingen.de
Human-Computer Interaction, University of Tübingen
Tübingen, Germany

Béla Umlauf
bela.umlauf@student.uni-tuebingen.de
University of Tübingen
Tübingen, Germany

Ard Kastrati
akastrati@ethz.ch
ETH Zurich
Zurich, Switzerland

Martyna Plomecka
martyna.plomecka@uzh.ch
Department of Psychology, University of Zurich
Zurich, Switzerland

William Schaefer
william.schaefer@utsa.edu
University of Texas at San Antonio
USA

Enkelejda Kasneci
enkelejda.kasneci@uni-tuebingen.de
Human-Computer Interaction, University of Tübingen
Tübingen, Germany

Zoya Bylinskii
bylinski@adobe.com
Adobe Research
USA

Figure 1: A gaze-based study design for assessing how users interact with a tutorial to accomplish a masking task in Adobe's Photoshop. Over the course of two tasks, we found competency of the task increased and reliance on the tutorial decreased. Eye movement differences were also indicators of better competency in the second task. Adobe stock photos featured ©fizkes, mahathir mohd yasin/EyeEm, Rachael Arnott – stock.adobe.com.

CCS CONCEPTS

• **Human-centered computing** → **Usability testing**; Human computer interaction (HCI); • **Applied computing** → *Interactive learning environments*; **Psychology**.

KEYWORDS

Eye Movements, Usability, Photoshop, Multimedia Learning

ETRA '22, June 8–11, 2022, Seattle, WA, USA

ACM ISBN 978-1-4503-9252-5/22/06.
https://doi.org/10.1145/3517031.3531634

ACM Reference Format:
Nora Castner, Béla Umlauf, Ard Kastrati, Martyna Plomecka, William Schaefer, Enkelejda Kasneci, and Zoya Bylinskii. 2022. A gaze-based study design to explore how competency evolves during a photo manipulation task. In *2022 Symposium on Eye Tracking Research and Applications (ETRA '22), June 8–11, 2022, Seattle, WA, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3517031.3531634

1 INTRODUCTION

A graphic design tool like Adobe's Photoshop provides a natural sandbox for exploring differences in gaze behavior as competency evolves, even in a single sitting [Lewien 2021]. The ability to help participants (i) excel in a specific task, (ii) by presenting a sequence of steps to complete it, and (iii) leveraging AI-driven tools to augment their abilities makes it possible to accomplish a naturalistic design task in a short time. Moreover, [Navarro et al. 2015] established that eye-tracking offers a unique insight to understanding a learner's experience and potential moments of confusion [Salminen

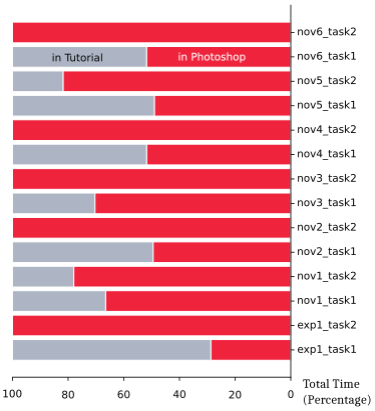


Figure 2: Percent time in tutorial and Photoshop across 6 novices and 1 expert.

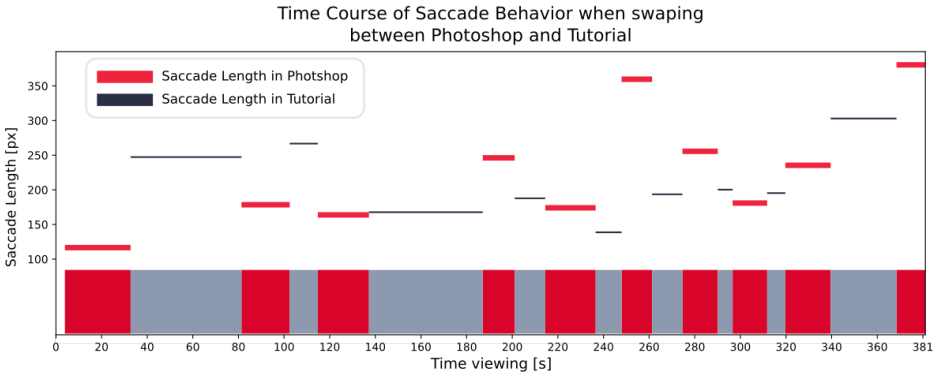
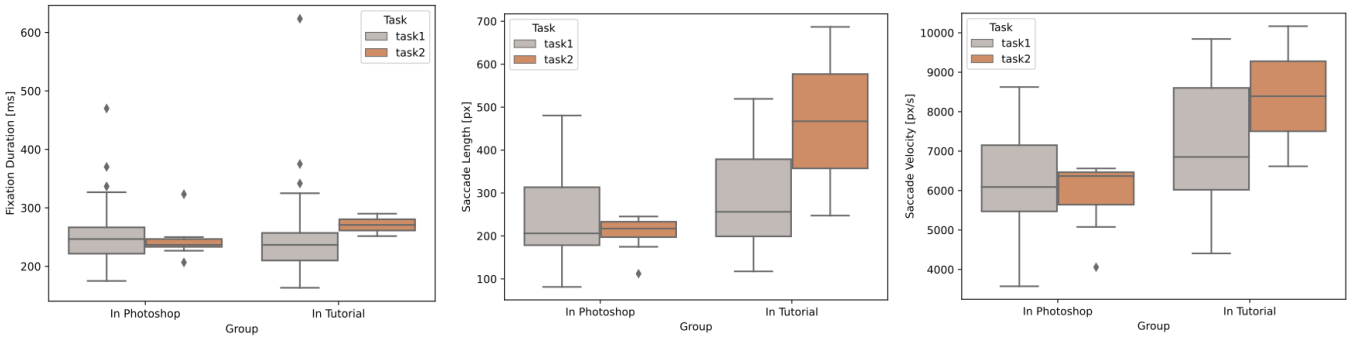


Figure 3: Time course of saccade length related to switching between tutorial (gray) and Photoshop (red).



(a) Fixation Duration [ms] **(b) Saccade Length [px]** **(c) Saccade Velocity [°/s]**

Figure 4: Gaze Behavior performing both masking tasks in Photoshop versus watching the tutorial.

et al. 2019]. We introduce a task setup optimized for analyzing gaze behavior during an image manipulation task in Photoshop. Seven participants completed a photo masking task in roughly three minutes, following a sequence of steps presented in a 2 minute video tutorial[1] that leverages automatic tools to simplify the task. Our pilot experiment demonstrates that having participants complete two such tasks in sequence already reveals differences in gaze behavior from the first to the second task. The study design allows participants to continuously toggle back to the tutorial and navigate it, providing opportunities to analyze when they get stuck and how they resolve their confusion. We show how a short experiment can already provide a rich space to explore how gaze relates to the evolution of competency during a task. Our study was motivated by how Adobe's Photoshop is taught in classrooms and our longer term goal is for the findings to be used in the benefit of educational programs [Sharma et al. 2020].

2 METHODS

Eye movements were recorded using the Tobii Pro Spectrum running at 300 Hz on a color-calibrated monitor with full HD resolution. We used the software *Titta*[2] in Python running the eye tracker and OpenCv to record the screen. Swapping between a full screen of either Photoshop or the video tutorial was done using the alt+tab key combination and these timestamps were extracted to map gaze to the respective window. Calibration was performed after watching the tutorial and between tasks. Raw gaze data was cleaned and event detection was performed using I-VT [3] with minimum fixation duration of 60 *ms* and a velocity threshold of 30 °/*s*.

3 RESULTS

Participants spent more time toggling between Photoshop (Figure 2, red) and the tutorial (gray) during the first task, spending roughly 52% of the time in Photoshop. Five of the total seven participants did not refer back to the tutorial at all during the second task.

[1] https://utsa.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=d633e6d9-7ecc-499f-b519-ae3c0158d5d4

[2] https://github.com/marcus-nystrom/Titta

[3] https://pypi.org/project/Perception-Engineers-Toolbox/

Gaze behavior exhibits differences between performing the masking task in Photoshop versus viewing the tutorial (Figure 4, left versus right part of each graph), particularly during the second task (brown bars). Fixation durations increased in the second task when viewing the tutorial (across the two participants that had to refer back to the tutorial for this task, Figure 4a). Larger saccade lengths (Figure 4b) and saccade velocities (Figure 4c) are observed when comparing the tutorial to masking in Photoshop in the second task (with a slight trend for the same pattern in the first task). There is also a trend (though not significant) for longer and faster saccades when viewing the tutorial during the second task compared to the first task. These behaviors could be indicative of participants more actively searching the tutorial (as opposed to more passively viewing it) during their second task, and more actively searching within the tutorial compared to within the Photoshop UI when performing the task. We have yet to analyze whether the larger saccades in these cases can be attributed to participants navigating between the video and video controls, or within the video itself.

Figure 3 shows an example of one participant toggling between Photoshop and the tutorial over the course of a masking task. Time spent on the tutorial decreases over time (gray bars at the bottom) while saccade lengths during these intervals remain steady around 150-250 pixels (y-axis), with the exception of the last saccade. In comparison, saccade lengths during task completion in Photoshop (red) are more variable between 100-400 pixels. It will be interesting to compare where saccades start and end within the recorded Photoshop UI (in the tutorial) compared to the actual Photoshop UI used during the masking task.

While this study was run as a pilot of our task and tutorial design, our next steps will be to run a larger set of participants, to compare novices and experts across both masking tasks and tutorial viewing.

4 DISCUSSION

The goal of this study was to introduce a task design that could be completed in a short sitting, while providing the ability to analyze differences in gaze behavior as competency evolves. This task mimics a self-guided learning scenario where students may toggle between tutorials (educational material) and the assignment at hand. Our early findings suggest that gaze behavior can be used to differentiate between the first and second time participants complete similar tasks. Our next steps will use these gaze measurements during the task to predict the likelihood that they would trigger the tutorial, as a signal for where confusion may be arising. This may help future implementations of design tools that trigger tutorials at the most helpful times for novice users, using gaze as a guide.

REFERENCES

Ryan Lewien. 2021. GazeHelp: exploring practical gaze-assisted interactions for graphic design tools. In *ACM Symposium on Eye Tracking Research and Applications*. 1–4.

Oscar Navarro, Ana Isabel Molina, Miguel Lacruz, and Manuel Ortega. 2015. Evaluation of multimedia educational materials using eye tracking. *Procedia-Social and Behavioral Sciences* 197 (2015), 2236–2243.

Joni Salminen, Mridul Nagpal, Haewoon Kwak, Jisun An, Soon-gyo Jung, and Bernard J Jansen. 2019. Confusion prediction from eye-tracking data: experiments with machine learning. In *Proceedings of the 9th International Conference on Information Systems and Technologies*. 1–9.

Kshitij Sharma, Michail Giannakos, and Pierre Dillenbourg. 2020. Eye-tracking and artificial intelligence to enhance motivation and learning. *Smart Learning Environments* 7, 1 (2020), 1–19.