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

NORA CASTNER and BÉLA UMLAUF, Human-Computer Interaction, University of Tübingen, Germany
ARD KASTRATI, ETH Zurich, Switzerland
MARTYNA PLOMECKA, Department of Psychology, University of Zurich, Switzerland
WILLIAM SCHAEFER, University of Texas at San Antonio, USA
ENKELEJDA KASNECI, Human-Computer Interaction, University of Tübingen, Germany
ZOYA BYLINSKII, Adobe Research, USA

Fig. 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 movements differences were also indicators of better competency in the second task.

CCS Concepts: • **Human-centered computing** → Human computer interaction (HCI); • **Applied computing** → **Psychology**.

Additional Key Words and Phrases: Eye Movements, Usability, Photoshop, Multimedia Learning

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 *Symposium on Eye Tracking Research and Applications (ETRA '22 Technical Abstracts), June 8–11, 2022, Seattle, Washington.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3379155.3391320

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 [1]. 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, [2] established that eye-tracking offers a unique insight to understanding a learner's experience and potential moments of confusion [3]. We introduce a task setup optimized for analyzing gaze behavior during an image manipulation task in Photoshop. Participants completed a photo masking task in rougly 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 Photoshop is taught in classrooms, and our longer term goal is for the findings to be used in the benefit of educational programs [4].

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 the gaze to either the respective window. Calibration was performed after watching the tutorial and before the second task. Raw gaze data was cleaned and event detection was performed using an implementation of the I-VT from the *Perception Engineer's Toolkit*[3] with minimum fixation duration of $60ms$ and a velocity threshold of $30°/s$.

Results Participants spent more time toggling between Photoshop (Figure 2, red) and the tutorial (gray) during the first task, spending roughly 52.55% of the time in Photoshop. Five participants did not

ETRA '22 Technical Abstracts, June 8–11, 2022, Seattle, Washington

[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/

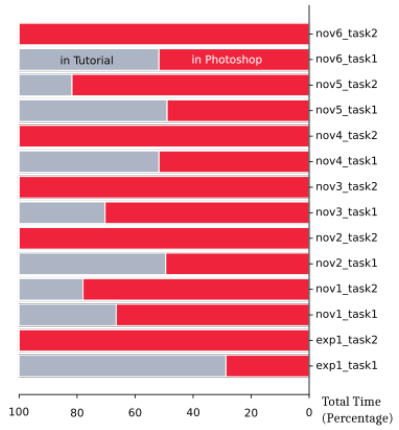


Fig. 2. Percent time in tutorial and Photoshop.

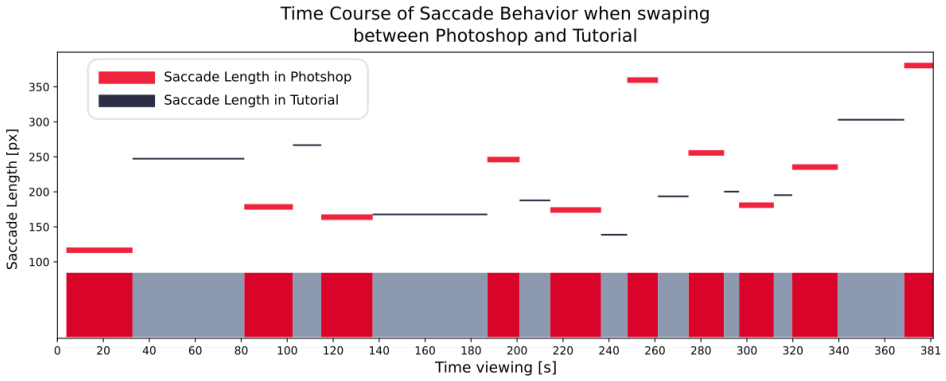
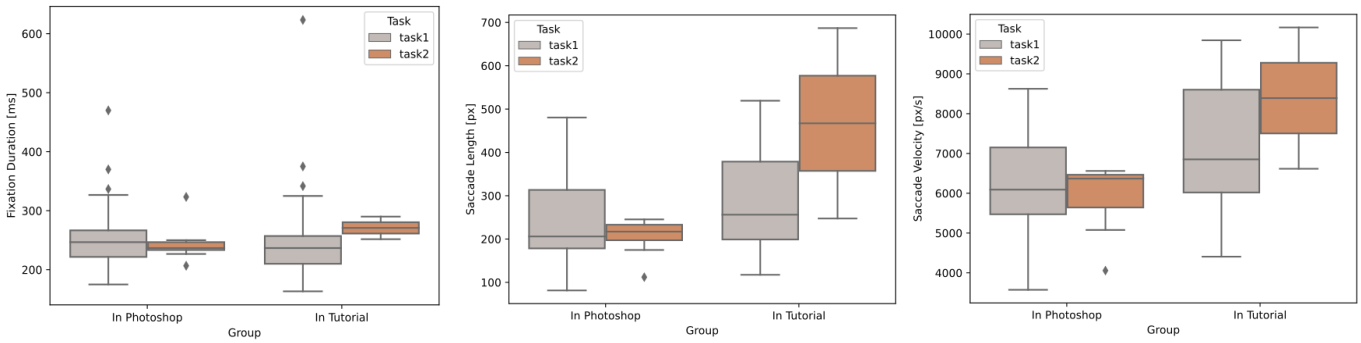


Fig. 3. Time course of saccade length related to switching between tutorial (gray) and Photoshop (red).



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

Fig. 4. Gaze Behavior performing both masking task in Photoshop versus watching the tutorial.

refer back to the tutorial during the second task. Gaze behavior shows clear differences between performing the masking task versus viewing the tutorial and, more interesting, between the first and second tasks (Figure 4). Fixation durations increased in the second task, most notably when viewing the tutorial. However, gaze duration showed little difference between viewing the tutorial and performing the task (Figure 4a). Saccade behavior (Figures 4b and 4c) exhibited the greatest differences, with larger lengths and faster velocities for the second task. Similarly, the tutorial evoked larger and faster saccade behavior in the second task. However, performing the first masking task evokes larger and faster saccades compared to viewing the tutorial. This behavior could be indicative of participants searching for the appropriate tools in Photoshop for the first time after viewing the tutorial. Figure 3 shows an example of how one participant toggles between Photoshop and the tutorial over the course of the task. Time spent on the tutorial decreases over time (gray bars at the bottom) while saccade lengths during these tutorial intervals remain steady around 150 pixels (y-axis). Saccade lengths during task completion in Photoshop (red) vary between 100 and 200 pixels.

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 trigger tutorials at the most helpful times for novice users, using gaze as a guide.

References

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

[2] 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.

[3] 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.

[4] 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.