

Peer Review

Review of: "Analyzing Audience Engagement in Static Versus Dynamic Media Content Using Eye-Tracking and Instagram Metrics Through the Lens of TPB and CLT"

Steven Bellman¹

1. Ehrenberg-Bass Institute, University of South Australia, Australia

This paper uses an ethically collected sample of data, which is good. It also provides some interesting evidence that the realeye.io application can be very inaccurate if participants are not properly calibrated. That application provides an interesting measure (K) that relates the time spent on fixations with the time spent on saccades. When video directors direct attention (<https://doi.org/10.1371/journal.pone.0142474>), fixation time is longer than saccade time ($K > 0$), whereas when people need to direct their attention themselves to make sense of static posters, saccade time is longer ($K < 0$).

However, the sample size is currently way too small for a study hoping to be published since Button et al. (2013, <https://doi.org/10.1038/nrn3475>). The design of the lab experiment is a within-participants design, comparing static stimuli (posters) with dynamic stimuli (videos). Assuming a small effect-size difference ($d = .3$) between these 2 stimuli, which is common for communication studies (Weber & Popova, 2012, <https://doi.org/10.1080/19312458.2012.703834>), G*Power suggests a sample size of 70 is needed to detect this difference at $p < .05$ with 80% power. Any significant differences reported in this paper may be spurious because the sample of 15 used did not have the power to detect them. The authors need to go back into the lab and get data from at least another 55 participants.

Another problem when comparing videos with static stimuli is that videos can have a long duration. A video lasting 60 seconds will demand attention and get more fixations and therefore more total fixation

duration than a static picture looked at for only 5 seconds (which may be enough time). The paper needs to report the durations of the stimuli and compare measures that control for duration, such as fixations per second. Using matching measures, there may be no difference between static and dynamic stimuli. Similarly, when comparing movie poster likes, it is important to use comparable rates, such as likes per views.

Currently, the paper uses eye-tracking metrics and likes as indicators of constructs in the Theory of Planned Behavior, citing Fielden et al. (2022) as a precedent. However, Fielden et al. used a validated survey to measure TPB constructs like attitude (8 adjective scales (*reassuring, unpleasant, embarrassing, unwise, important, worrying, worthwhile, and healthy*)). The paper needs to delete this analysis, as it is not valid. The same goes for the use of eye-tracking measures like fixation count and fixation duration as measures of cognitive load. These need validation from prior studies and triangulation (e.g., with self-reported measures of cognitive load) from other measures collected in the lab.

I think the current version of the paper needs a lot of work to be publishable. I hope the authors find these comments useful when revising the paper. I wish the authors the best of luck with their future research.

Declarations

Potential competing interests: No potential competing interests to declare.