# Eye Tracking in Emotional Design Research: What Are Its Limitations?

**Cyril Brom**
Faculty of Mathematics and Physics, Charles University in Prague, Czech Republic
brom@ksvi.mff.cuni.cz

**Tereza Stárková**
Faculty of Arts, Faculty of Mathematics and Physics
Charles University in Prague
Czech Republic

**Jiří Lukavský**
Institute of Psychology
Academy of Sciences of the Czech Republic
Czech Republic

**Ondřej Javora**
Film and TV School
Academy of Performing Arts in Prague
Czech Republic

**Edita Bromová**
Faculty of Mathematics and Physics
Charles University in Prague
Czech Republic


*NordiCHI '16*, October 23 - 27, 2016, Gothenburg, Sweden

ACM 978-1-4503-4763-1/16/10…$15.00
DOI: http://dx.doi.org/10.1145/2971485.2996749

## Abstract
Adding anthropomorphic graphical elements to computerized learning materials has been shown to enhance learning. We have prepared an experimental study, in which we will investigate if learning is enhanced because of the elevated positive affect of learners or due to the attention-capturing effect of the anthropomorphisms. Here, we present a pilot study asking if eye tracking might help in answering this question. Participants ($N$ = 23) studied from one of three versions (schematic, enhanced black-and-white, enhanced with color) of instructional slides about the influenza virus for about 10 minutes; with eye tracking. Their prior and, especially, post hoc negative affect was elevated, and positive affect decreased compared to participants from a similar experiment without an eye tracker ($N$ = 37). One-third of participants reported higher nervousness due to the eye tracker. The attention-capturing effect was possible to measure; but with limitations. The implications for emotional design studies with eye tracking are discussed.

## Author Keywords
Emotional design; anthropomorphisms; positive affect; negative affect; eye tracking; eyes; slides.

## ACM Classification Keywords
H.5.2. User interfaces: Evaluation/methodology.

## Introduction

Multimedia learning refers to learning from any material combining text and pictures [6], including textbooks, computerized presentations and animations or videogames. So-called multimedia learning principles (e.g. [3, 6]) show how to design these materials so that they can be cognitively processed by learners in an effective way. However, there is very limited knowledge on how to improve multimedia learning materials that have been already optimized from the cognitive perspective so that they are also affectively appealing and motivating and thus further enhance learning. These improvements are sought after in so-called *emotional design research* (e.g., [9, 12]). Our work is situated within this milieu and asks the following research questions:

1. if replacing schematic visual elements in a computerized presentation with anthropomorphisms with expressive eyes (see Figure 1):
   a) elevates the positive affect [13] of learners;
   b) improves learning outcomes;
   c) has attention-capturing effects as measured by eye tracking;
2. if learning outcomes are predicted by
   a) affective-motivational variables;
   b) gaze patterns;
   c) affective-motivational variables after the effect of gaze patterns is accounted for.

This paper presents a pilot study for this research, which (tentatively) addresses Questions 1a and 1c.

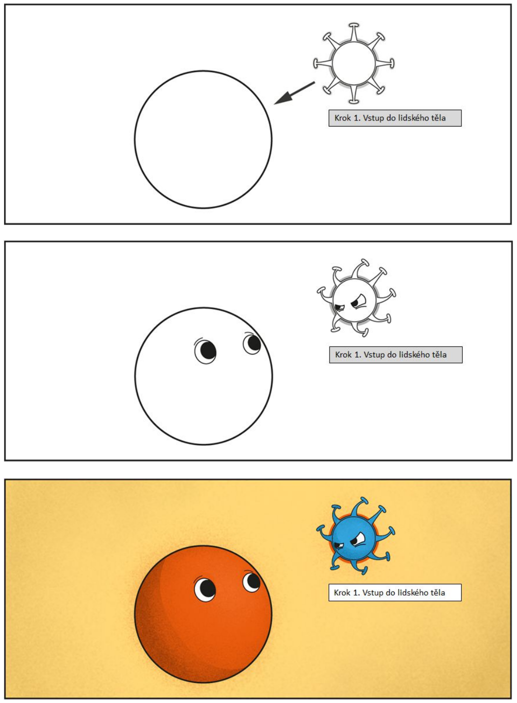


**Figure 1.** Graphics from the three versions of the slides compared in this experiment. Inspired by [7]. ©Ondřej Javora

## Background

*Opposing Effects of Affective Manipulations*

In the emotional design research field, there is a consensus that a manipulation having the ability to elevate positive affect (or motivation, interest, positive activating emotions) can have two effects with opposing impact on quality of learning (cf. [12]). First, it can activate the learner, increasing the effectivity of cognitive processing (improving learning). Second, it can consume cognitive resources (hindering learning), because a) it has to be processed in the first place, which requires cognitive resources, and b) it can lead to losing focus. This idea is theoretically underpinned, for instance, by the Cognitive-Affective Theory of Learning with Media [6]. It is unknown for which manipulations (and what types of learners) "affective" benefits outweigh disadvantages.

*Anthropomorphisms in Multimedia Learning*

Research on emotional design is in its infancy [9, 12]. The most promising manipulation found so far is adding anthropomorphisms (combined with warm, bright color or gray-scale) to brief animations or slides. Although the data are not always clear-cut, this manipulation tends to elevate affective-motivational states [10, 12] and improve learning [1, 7, 10, 12]. It is not clear if the positive effect of anthropomorphisms is of affective origin (e.g., they can also have signaling effects, which improve the way learners process the pictures). Only one study on the effect of anthropomorphisms has used an eye tracker [8]: it found attention-capturing effects of expressive eyes but not of simpler geometric eyes. It also reported no improvements to positive affect and mixed results concerning learning outcomes.

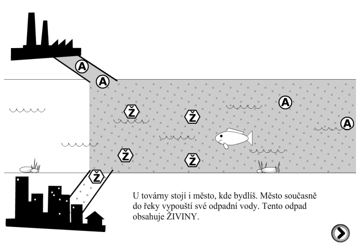


**Figure 2.** A screenshot of the animation used for comparison.

*This Study*

The aim of our whole study is to expand present knowledge on the usage of anthropomorphisms in multimedia learning materials by investigating the attention-capturing effects of eyes as a mechanism (partly) responsible for improved learning outcomes (cf. Research Question 2c from Introduction). In the current pilot study, we focus on Questions 1a and 1c. Given the known signaling effects of eyes [4], we reason that because of them, learners' processing of pictures will be improved. We thus refine our Research Question 1c by hypothesizing that:

i. expressive eyes will capture the learner's attention (i.e., learners will inspect the element with eyes for a longer period of time);
ii. learners will follow eye gaze (and thus then process the adjacent element, e.g., virus → cell on Fig. 1).

We have used learning materials by Mayer and Estrella [7] and optimized them for eye tracker usage (e.g., by normalizing the size of elements). The original materials included an anthropomorphic version with colors and a schematic version. We also added an anthropomorphic black-and-white version (Fig. 1) and we included more sophisticated instruments for measuring affective variables (not used in [7]).

Based on previous results, small to medium effect sizes for between-condition differences in affective variables can be expected. We cannot hope to detect such differences with the pilot's sample size. However, as a manipulation check (Question 1a), we ask if the affective variables will be in the same ranges compared to the affective states for a comparable sample engaged in a similar learning experience; namely, studying the functioning of a wastewater treatment plant from a brief animation (cf. [2], Fig. 2).

## Method

*Participants*

Undergraduate students of psychology – special education (17 females, 2 males; $M_{age}$ = 21.8±2.8 [±SD]) and undergraduate computer science students

**Eye Tracking Details**

**Model**: EyeFollower eye-tracking system (LC Technologies, Inc., USA), 120 Hz, 24-in screen, 75 cm dist.

**Areas of Interest (AOIs):** Three different levels of detail: 1) whole textual vs. pictorial part; 2) individual elements (e.g. cell, virus), 3) sub-elements (e.g. eyes of the virus vs. the rest of the virus, Figure 3).

**Analysis:** Each valid gaze sample (94.1%) assigned to a corresponding AOI. The dwell time analyzed for the critical AOIs and gaze transitions (where one's gaze lands next after visiting a critical AOI). Only transitions with no other transition in 50 ms considered.

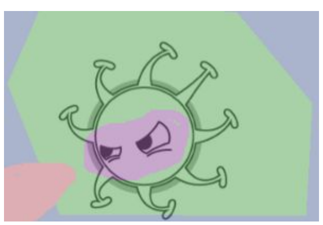

**Fig. 3.** AOIs example: the eyes (pink) vs. the virus (green) vs. a different object (red). ©O. Javora

(4 males; $M_{age}$ = 21.5±1.7) participated. None of them had participated before in an eye tracking experiment. For the eye tracking analysis, we checked the number of invalid gaze samples and fixations beyond the monitor and discarded data from six participants (with <80% valid samples). The mean proportion of valid samples in retained data was 94.1%±2.7.

*Materials, Procedure*

Participants were informed that they will study (from a few slides) how the influenza virus attacks the human body, that they would be tested on that knowledge thereafter and an eye tracker would be used as part of the research. After filling in a consent form, a demographic questionnaire, and a questionnaire on initial interest and anxiety (5 (3) items with a 7-point Likert scale; see [2]), the eye tracker was calibrated. Participants then received one of the slides versions (Fig. 1). The slides were self-paced (2 introductory textual slides; 9 text-and-picture slides: each with 1-5 elements with expressive eyes and one textual inset; 1 final textual slide; $M_{time\text{-}to\text{-}complete}$ = 10.6±4.1 min). The text was placed above the picture (650 words in total). After the learning part, the PANAS [5] and the Flow short scale (FSS) [11] were administered. The PANAS measured generalized positive and negative affect (each with 10 items with a 5-point Likert scale) and the FSS measured flow levels (10 items, 7-point Likert scale). Next, participants received an evaluation questionnaire irrelevant to the present purpose and knowledge tests (we calibrated these tests during this pilot). Finally, participants were debriefed.

*Sample and Materials Used for Comparison*

We used different undergraduate students of psychology – special education (20 females, 1 male; $M_{age}$ = 21.2±1.4) and computer science (15 males, 1 female; $M_{age}$ = 23.1±2.1) for comparison. These persons studied (in a different experiment [2]) from a black-and-white animation (Fig. 2) about the functioning of a waste wastewater treatment plant for roughly 6 minutes. The procedure and materials were comparable to the present pilot study, except for the eye tracker usage.

## Results and Discussion

*Do eyes capture attention?*

Learners spent more time on texts compared to pictures, regardless of the condition (331±132 s vs. 77±41 s). This is not surprising (e.g., cf. [5]). In the debriefing, learners clearly indicated their awareness of anthropomorphisms (in the two respective conditions). Descriptively, eyes captured more attention compared to the schematic version (19±14 s vs. 15±11 s; Figure 4), but a smaller amount of time was spent on eyes overall compared to the rest of the anthropomorphized elements (19±14 s vs. 26±12 s). Across all conditions, fixations on the virus were more frequent compared to fixations on the cell (27±16 s vs. 18±14 s); probably due to more complex shape of the virus (the mean area of virus AOIs represented a third of mean cell AOIs).

In the most detailed AOI analysis (i.e., level 3), the total number of transitions from eyes to the element that the eyes were looking at (e.g., eyes of the virus → cell) was negligible in all conditions (10±12). However, transitions from eyes to the rest of the element (e.g., eyes of the virus → virus) occurred substantially more often (75±49). Analysis of transitions on a coarser level (level 2; e.g. virus → cell) may thus provide a better measure (77±42); with sufficient variability to test further hypotheses and robustness against noise.

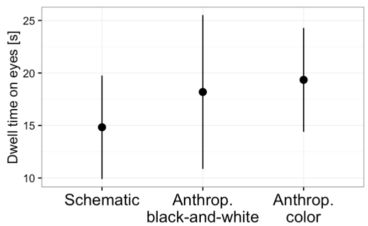


**Figure 4.** Time spent on eyes by condition (mean and SE).

| | Waste water | This pilot study |
|---|---|---|
| Prior interest | 21.0 \| 20 (4.9) | 23.5 \| 24 (4.4) + |
| Prior anxiety | 10.8 \| 11 (2.5) | 12.8 \| 13 (2.7) ** |
| Prior neg. aff. | 14.7 \| 14 (3.9) | 18.0 \| 18 (6.6) * |
| Prior pos. aff. | 28.6 \| 28 (6.0) | 27.5 \| 27 (6.2) n.s. |
| Post hoc neg. aff. | 12.1 \| 12 (2.4) | 17.0 \| 14 (7.2) ** |
| Post hoc pos. aff. | 31.8 \| 31 (6.5) | 27.9 \| 27 (7.7) + |
| Post hoc flow | 56.8 \| 59 (7.3) | 44.9 \| 47 (9.8) *** |

**Table 1.** Affective variables: means | medians (standard deviations). Significance tested by Mann-Whitney test ($+p < .1$; $*p < .05$; $p** < .01$; $p*** < .001$).

Taken together, these data indicate that in the main study we are planning, the attention-capturing effect of eyes could indeed be found (Hypothesis i). To inspect its postulated consequences, i.e., gaze cueing (Hypothesis ii), it will be probably necessary to analyze transitions between whole anthropomorphized elements rather than just from eyes. The same will likely be needed in any similar study.

*Do anthropomorphisms elevate positive affect?*

An unexpected result was found when comparing current affective data to data from the wastewater treatment experiment (Table 1). Despite comparably higher initial interest among this pilot's participants, their initial anxiety, prior and, especially, post hoc negative affect were also elevated. Participants' post hoc positive affect and flow levels were substantially lower. Part of these differences could be explained by a) the two participant samples used for comparison being slightly different and b) the materials and the topic being different (e.g., slides vs. animation). However, one-third of participants complained in the debriefing that they were distracted and stressed by the eye tracker. This model's camera made some noise when searching for eyes (This happened quite often, because during the learning experience participants did not fixate on the image all of the time. They also often looked outside the screen when they were mentally processing the learning message.). One-fourth of participants complained about noise, but this could explain post hoc differences, not prior differences. Indeed, five participants suggested that they were concerned (prior to the experiment) with the fact that their eyes would be monitored (a kind of "Big Brother" effect). The distribution of negative affect (Fig. 5) suggests that this is unlikely an issue for all.

Taken together, these findings warn us that the affective state of participants may also be influenced by the eye tracking research method per se. This means that for the main study, either participants influenced by the presence of the eye tracker will have to be excluded, or different participants will have to be recruited to investigate Questions 1a, 1b and 2a vs. 1c and 2b (leaving Question 2c out). We have opted for the latter option. The main study will continue during the rest of 2016 and the beginning of 2017.

## Conclusion

This pilot study investigated if the eyes of anthropomorphized elements in instructional slides can capture attention and elevate positive affect when attention-capturing effects are measured using an eye tracker. The results suggested that a) attention-capturing effects (if any) can be found in the main study using the present research method, b) the majority of transitions from eyes (or the complementary area in the schematic version of the slides) ends in the remaining part of the element with the eyes, and c) the presence of the eye tracker may distract or even be stressful for some participants.

To our knowledge, this is the first study that attempts to analyze out-going transitions from eyes in the context of emotional design research. The methodological implication for future studies (including our own) is that analysis of transitions from whole elements could be needed. This is because there may be too few transitions leading from the eyes outside to the element containing the eyes for a meaningful analysis.

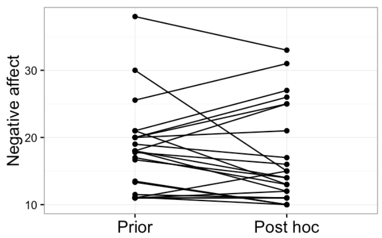


**Figure 5.** Prior and post hoc negative affect for each participant.

A wider methodological implication is that using an eye tracker in studies where the affective variables are of key interest may be problematic. This is because an eye tracker's presence may influence the affective state of some participants. Due to the preliminary nature of our findings, it is obviously still a question to what extent our results would generalize. It would be informative if other researchers check if participants in their studies were influenced the same as in our pilot study.


## Acknowledgements

This research was partially funded by the research project nr. 15-14715S supported by the Czech Grant Science Foundation (GA ČR).


## References


[1] Brom C., Hannemann, T., Starkova, T., Bromova, E., & Dechterenko, F. Anthropomorfic faces and funny graphics in an instructional animation may improve superficial rather than deep learning: A quasi-experimental study. In *Proc. of European Conference on e-Learning* (2016). (accepted)

[2] Brom, C., Hannemann, T., Stárková T., Bromová, E., & Děchtěrenko, F. The role of cultural background in the personalization principle: Experiments with Czech learners (2016). (submitted)

[3] Clark, R.C. and Mayer, R.E. *E-learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning*. John Wiley & Sons 2011.

[4] Frischen, A., Bayliss, A.P. and Tipper, S. P. Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological bulletin*, 133, 4 (2007) 694-724.

[5] Johnson, C.I., & Mayer, R.E. An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied*, 18, 2 (2012), 178-191.

[6] Mayer, R. *Multimedia Learning*. Cambridge University Press, 2009.

[7] Mayer, R. E. and Estrella, G. Benefits of emotional design in multimedia instruction, *Learning and Instruction*, 33 (2014), 12-18.

[8] Park, B., Knörzer, L., Plass, J. L. and Brünken, R. Emotional design and positive emotions in multimedia learning: An eyetracking study on the use of anthropomorphisms. *Computers & Education*, 86 (2015), 30-42.

[9] Plass, J.L. and Kaplan, U. Emotional Design in Digital Media for Learning. In *Emotions, Technology, Design, and Learning* (2015), 131–161.

[10] Plass, J.L., Heidig, S., Hayward, E.O., Homer, B.D. and Um, E. Emotional design in multimedia learning: Effects of shape and color on affect and learning, *Learning and Instruction*, 29 (2014), 128-140.

[11] Rheinberg, F., Vollmeyer, R. and Engeser, S. Die Erfassung des Flow-Erlebens [in German]. In *Diagnostik von Motivation und Selbstkonzept*, Hogrefe 2003, 261-279.

[12] Um, E.R., Plass, J.L., Hayward, E.O. and Homer, B.D. Emotional Design in Multimedia Learning, *Journal of Educational Psychology*, 104, 2 (2012) 485-498.

[13] Watson, D., Clark, L.A. and Tellegen, A. Development and validation of brief measures of positive and negative affect: the PANAS scales, *Journal of Personality and Social Psychology*, 54, 6 (1988), 1063-1070.