Exploring Eye-Tracking Data for Detection of Mind-wandering on Web Tasks

Jacek Gwizdka

School of Information, University of Texas at Austin, TX, USA
neurois2018@gwizdka.com

Abstract
Mind-wandering (MW) is a phenomenon that affects most of us; it affects our interactions with information systems. Yet the literature on its effects on human-computer interaction is only scant. This research aims to contribute to establishing eye-tracking measures that could be used to detect periods of MW while a user is engaged in interaction with online information. We conducted a lab study (N=30) and present an exploratory analysis of eye-tracking data with a focus on finding differences between periods of MW and not-MW. We found 12 eye tracking measures that were significantly different between periods of MW and not-MW. We also show promising classification results of the same variables. Our results indicate plausibility of using eye-tracking data to infer periods of MW.

Keywords: mind-wandering, mindless reading, eye-tracking, pupillometry.

1 Introduction

Most people have experienced mental state when their mind has wandered. This phenomenon is quite common and many people can remember when, for example, their reading did not result in any meaningful understanding of the text. In this case, a wandering mind can be a harmful thing as it makes us less efficient, prone to errors and to making incorrect decisions. If an information system were able to detect when a person's mind is wandering, it could offer an intervention. For example, if it detects that an e-commerce website user has spent a significant time MW while reviewing purchase options, the system could ask for additional verification before the purchase is made.

The goal of this project is to establish eye-tracking based measures of MW that could be used to detect periods of MW while a user is engaged in interaction with online information. We present an exploratory analysis of eye-tracking data (including pupillometry) with a focus on finding differences between periods of task-related and task-unrelated thoughts.

Gwizdka, J. (2019). Exploring Eye-Tracking Data for Detection of Mind-Wandering on Web Tasks. In F. D. Davis, R. Riedl, J. vom Brocke, P.-M. Léger, & A. B. Randolph (Eds.), Information Systems and Neuroscience (pp. 47–55). Springer International Publishing. https://doi.org/10.1007/978-3-030-01087-4_6

This is last author's manuscript version.

2 Related Work

While it is known that MW can have positive effects on human thought processes and, in particular, on creativity [1], MW is typically detrimental to the tasks that require focused attention. MW has been shown to negatively affect reading [2, 3], and the ability to resolve conflicts in displayed information [4], because the executive function is impaired. These are just two examples of negative influence of MW on user interaction with IS. MW has received increased attention from cognitive scientists and psychologists in the last decade [1, 3, 5], but we are only beginning to understand these processes.

Interestingly, it has been demonstrated that MW while reading is related to changes in eye-fixation patterns. For example, fixation durations were found to be longer and less affected by lexical and linguistic variables, while eye movements were found more erratic during MW periods than during reading [2]. Pupil dilation was found to change more spontaneously during MW periods [6, 7]. Results from these studies suggest that eye movements during MW are controlled by different cognitive processes than during normal reading. People may engage in internally focused cognitive tasks, which, presumably, are associated with different cognitive processes than externally focused tasks and unfocused MW. Recent work compared eye movements during goal-directed internally focused cognitive task and reading task. Eye behavior during the former was different and characterized, among others, by more and longer blinks, fewer microsaccades, more and shorter fixations, more saccades and saccades of higher amplitude [8]. These results suggest that some aspects of eye behavior may be coupled with internally generated information and related internal cognitive processes. We will come back to this in the Discussion section.

While MW is an important phenomenon with potentially significant explanatory power for human interaction with IS, research in this area outside psychology is still rather scant. Notable exceptions include, a theoretical model of MW in a technological setting proposed in a doctoral dissertation, with a specific goal to better understand costs and benefits of this phenomenon on technology users [9, 10]; a person-independent detection of MW based on eye-tracking data proposed by a group of computer scientists [11]; and the use of eye-tracking data and web cams in a large scale detection of MW in an education online setting [12].

Encouraged by previous research that showed relationship between episodes of MW and eye-tracking data, we aimed to 1) use more realistic stimuli than in psychology research (e.g., [6, 7]), 2) perform analysis without considering text characteristics (in contrast to [2] and to local features reported in [11]), and 3) examine differences in eye-tracking variables between periods of MW and nMW (not reported in [11], but reported in their second paper [13]).

Our research questions are as follows:

RQ1. Which eye-tracking measures differ significantly between periods of MW and nMW?

RQ2. Can eye-tracking measures be used to classify periods of MW and nMW?

3 Method

We conducted an eye-tracking lab experiment (N=30, 20 females) in Information eXperience (IX) lab in the School of Information at University of Texas at Austin. Eye tracking data was collected using remote eye-tracker Tobii TX-300. The experiment was approved by IRB. Each lab session typically lasted 30 minutes. At the completion of a session, each participant received $12.

3.1 Procedure and Materials.

Participants were pre-screened for their native or near-native level of English, and for normal to corrected-to-normal eyesight. Each participant filled out background questionnaire, performed a training task, and three online reading tasks shown in randomized order. After each task participants answered comprehension questions. These questions were included to provide motivation for attentive reading. The task design followed a simulated work-task approach [14], where tasks are presented with reasons for their performance. In our study, participants were informed that they needed to read three articles in order to prepare for a course “Technology and Society” they were taking. The articles were taken from the UBC-Hampton Reading Comprehension Test Suite [15]; their sources are listed in Table 1. Each text was presented on several web pages and was displayed in black Arial font on white background. The pages were designed to show about the same number of text lines on the page and thus each screen had about the same luminescence. Text lines height was uniform at 27px, which corresponds to 0.45° of visual angle and is approximately equal to the eye-tracker's accuracy reported by the manufacturer as 0.4°-0.5°.

Table 1. Articles for online reading tasks.

1–*A quick overview of digital activism. A blog post by Curiouscatherine (2011)*
2–*Taking the slack out of slacktivism. A popular press article published by Radio Free Europe, 2011.02.17*
3–*'Free the spectrum!' Activist encounters with old and new media technology. A journal article by Dunbar-Hester, C., published in New Media & Society*

To capture incidents of MW, participants were periodically probed [16] and asked to indicate whether they were reading or MW [17]. We used a visual probe [18] – a pop-up window with two response buttons (Fig 1), which was displayed at random times controlled to be between 40 and 60 seconds and shown for 12 to 16 seconds at several pseudo-randomly selected locations on screen.

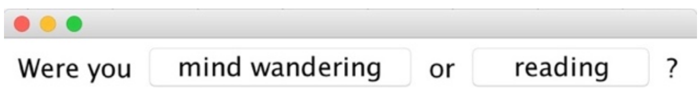

Fig 1. Pop-up window with the mind-wandering probe.

3.2 Variables

Independent variable was MW state with two levels: MW or nMW (i.e. reading). Dependent variables were obtained from eye-tracking data for two 5-second-long epochs that started respectively 5 and 10 seconds before the response to the probe. Eye-tracking data was cleaned by removing bad quality fixations (as marked by Tobii). Data from epochs with very few fixations (<3 standard deviations (std) below the mean) was discarded (<5% of epoch data) as it indicated epochs with many missing data points. We used eight types of variables: fixation count, regressions count, fixation duration, saccade duration, length, velocity, angle (four categories), and relative pupil change. For fixation duration and numerical saccade measures we calculated: mean, std, total, min and max values. For relative pupil change, which we calculated separately for each user and then denoised it with cubic interpolation of missing data (e.g., blinks), we calculated mean and std. Saccade angle was categorized to indicate (approximately) eye movement 1) forward (-10°; 10°), 2) backward (170°; 180°) in the same text line. Angles outside these ranges were categorized as 3) forward or 4) backward above or below the text line. This process yielded twenty-seven (27) variables.

4 Data Analysis and Results

Two participants have reported no periods of MW. Following prior work [6], we treated them as "outliers" and removed their data. Thus, we report data from 28 participants. The obtained proportion of MW responses (27.3%) matches the expectations [11, 13]. Response times to the probe were significantly longer for MW as compared with nMW periods (**Table 2**). This supports participants' correct self-classification of their internal thought processes.

Table 2. Responses to the probe.

Mind-wandering	Count	%	Response time [ms] mean (sd)
YES	115	27.3	3159 (1192)
NO (reading)	306	72.7	2879 (933)
Total	421	100	M-WU: z=2.2, p=0.03

We performed inferential statistics and classification. In inferential statistics, due to not-normal distribution of variables and lack or homogeneity of variance, we used non-parametric Mann-Whitney U test (M-WU). Given the exploratory nature of our research, we conducted individual M-WU tests on each variable (**Table 3**).

We run classifications using Weka 3.8 [19]. Due to the imbalanced number of samples between MW and nMW classes, we used two data sampling methods to improve the balance, 1) synthetic sample generation SMOTE [20] and 2) random sample generation with replacement. We present best classification results obtained by applying random forest classifier with 10-fold cross-validation [21] (**Table 5**) and the best features for each classifier (**Table 4**). These features overlap with significant results in (**Table 3**).

Table 3. Descriptive statistics and Mann-Whitney U tests for significantly different variables.

Epoch	Variable	# data samples (read/MW)	MW Mean(sd)	nMW Mean(sd)	M-W U
5s	total_fixation_duration	306/115	3792(761)	3915(764)	z=2.44; p=0.015
10s	avg_fixation_duration	288/110	254(87)	238(36)	z=--2.02; p=0.043
5s	fixation_count	306/115	16(3)	16.5(3.5)	z=2.02; p=0.043
10s	- „ -	288/110	16(3)	16.8 (3.2)	z=2.46; p=0.014
10s	tot_saccade_len	288/110	2072(909)	2220(847)	z=2.2; p=0.027
10s	max_saccade_len	288/110	710(355)	779(342)	z=1.97; p=0.049
5s	max_saccade_dur	306/115	362(307)	301(357)	z=-3.18; p=0.0015
10s	- „ -	288/110	324(311)	298(344)	z=-2.78; p=0.0055
5s	avg_saccade_dur	306/115	76(68)	71(90)	z=-2.36; p=0.019
10s	- „ -	288/110	68(63)	64(67)	z=-2.74; p=0.0061
5s	std_saccade_dur	306/115	105(109)	91(154)	z=-3.1; p=0.002
10s	- „ -	288/110	98(132)	85(130)	z=-2.98; p=0.0028
10s	min_saccade_vel	288/110	.6(.68)	.63(2.3)	z=3.02; p=0.0025
10s	angle_cat_bck_count	288/110	3.2(1.6)	3.7(1.9)	z=2.17; p=0.03
5s	avg_pupil_change	306/115	-.015(.06)	-.0042(.05)	z=1.93; p=0.0535
10s	- „ -	288/110	-.015(.06)	-.0019(.05)	z=2.55; p=0.011
5s	std_pupil_change	302/109	.034(.015)	.027(.014)	z=--4.3; p<0.0001

Table 4. Classification results – best features

Epoch	Data sampling	**Best features** (in the order of weights from Information Gain Ranking Filter)
5s	intact	std_pupil_change , std_saccade_dur
	SMOTE (synthetic)	angle_cat_fwd_ud_count, regression_count, std_pupil_change, std_saccade_dur, max_saccade_dur, min_saccade_vel
	Random with replacement	std_pupil_change, avg_saccade_dur, std_saccade_dur, max_saccade_dur, min_saccade_vel, angle_cat_fwd_count
10s	intact	max_saccade_dur, std_saccade_dur, avg_saccade_dur
	SMOTE (synthetic)	std_saccade_dur, max_saccade_dur, avg_saccade_dur, avg_pupil_change
	Random with replacement	std_saccade_dur, avg_pupil_change, avg_saccade_dur, angle_cat_bck_count, max_saccade_dur, angle_cat_bck_ud_count, std_pupil_change

5 Discussion

Responding to *RQ1*, we found that 12 eye-tracking measures (44% of all measures considered) significantly differed between periods of MW and nMW – seven in 5s epoch and ten in 10s epoch (in that, five measures overlapped in both epochs). These results, taken together with the confirmatory answer to *RQ2*, i.e. the reasonably promising classification results, indicate plausibility of using eye-tracking data to infer periods of MW, at least on the tasks similar to ours.

Our results generally match the previous research. For example, similarly as in [2], we found that average fixation duration (avg_fixation_duration) tended to be longer in MW periods (in 10s epochs). We also found a higher variability of changes in pupil dilation (higher avg_pupil_change and std_pupil_change) in MW periods. This is similar to results presented in [6, 7], where the authors reported more spontaneous changes in pupil dilation during MW periods. [13] found a longer minimum saccade duration in MW periods, while we found mean and maximum saccade duration to be longer during MW, as well as a higher saccade duration variability (std_saccade_dur) during MW.

Several findings from our study are in some contrast to [8]. We found fewer fixations (in both 5s and 10s epochs) and longer average fixation duration (in 10s epochs) in MW periods, while [8] reported more and shorter fixations on their goal-directed internally focused cognitive task. This suggests that internal cognitive processes during MW periods caught in our study are different from cognitive processes during internally focused cognition described in [8]. It further suggests that the differences in cognitive processes associated with internally focused cognition, MW and reading are reflected in eye behavior and thus can be discerned from eye-tracking measures.

Table 5. Classification results (Random Forest with 10-fold cross validation).

Epoch	Data sampling	Samples nMW/MW	Accuracy [%]	ROC [%]	F-measure [%]	F-measure: for MW class [%]
5s	intact (no sampling)	305/115	73.2	62.7	68.3	28.0
	SMOTE (synthetic)	306/230	77.4	84.0	77.3	72.8
	Random with replacement	210/210	87.6	96.2	87.6	87.9
10s	intact (no sampling)	288/110	70.9	59.3	64.3	17.1
	SMOTE (synthetic)	288/220	76.0	83.2	75.8	71.1
	Random with replacement	199/199	89.0	96.3	88.9	89.1

Compared with [10, 12], accuracy of our classifications is higher (73%-89% in our work vs. 59%-72% and 52%-74% in their first and second work, respectively). This difference may be due to our use of different classification algorithm and a somewhat different set of features. However, our low values of F-measures for MW class (**Table 5**) indicate poor classification performance for this class when no resampling was used. This points to the need for more data samples from each user.

Contrary to prior work [2, 10], we did not find a significant difference in regression counts. Although it was marginally significant at p=.098 for 5s epoch, the difference was in the direction opposite to the expected, that is, we found fewer regressions in MW than in nWM periods. The same unexpected relationship was found in the related measure angle_cat_bck_count, which was significantly different in 10s epoch. We also have not found a higher number of line crossing saccades during MW (reported in [13]). We don't have yet a good explanation for these findings.

5.1 Limitations

Limitations of our work include unbalanced number of data samples from MW and nMW segments. This is expected and is due to typical frequency of MW occurrences and suggests the need to collect more data before classification algorithms are trained. We also plan to use a wider variety of epoch lengths and in the future studies, use different tasks.

6 Conclusion

We believe that MW is a phenomenon that will grow in importance and will be more widely studied in the context of human interaction with information systems. A broader impact of implicit detection of MW lies in its potential applicability to e-commerce and e-learning systems, where upon detection of MW an intervention could be offered to a customer or learner.

7 Acknowledgements

This project was supported, in part, by the Temple Teaching Fellowship 2016-2017. We thank Ms. Xueshu Chen, who was a Graduate Research Assistant, for her contributions to this project.

References

1. Schooler, J.W., Mrazek, M.D., Franklin, M.S., Baird, B., Mooneyham, B.W., Zedelius, C., Broadway, J.M.: The Middle Way: Finding the Balance between Mindfulness and Mind-Wandering. In: Brian H. Ross (ed.) Psychology of Learning and Motivation. pp. 1–33. Academic Press (2014).
2. Reichle, E.D., Reineberg, A.E., Schooler, J.W.: Eye Movements During Mindless Reading. Psychol. Sci. 21, 1300–1310 (2010).
3. Smallwood, J.: Mind-wandering While Reading: Attentional Decoupling, Mindless Reading and the Cascade Model of Inattention. Lang. Linguist. Compass. 5, 63–77 (2011).
4. Choi, H., Nam, C.S., Feng, J.: A Wandering Mind Cannot Resolve Conflicts in Displayed Information. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 59, 1397–1401 (2015).
5. Smallwood, J.: Distinguishing how from why the mind wanders: A process–occurrence framework for self-generated mental activity. Psychol. Bull. 139, 519–535 (2013).
6. Franklin, M.S., Broadway, J.M., Mrazek, M.D., Smallwood, J., Schooler, J.W.: Window to the wandering mind: Pupillometry of spontaneous thought while reading. Q. J. Exp. Psychol. 66, 2289–2294 (2013).
7. Smallwood, J., Brown, K.S., Tipper, C., Giesbrecht, B., Franklin, M.S., Mrazek, M.D., Carlson, J.M., Schooler, J.W.: Pupillometric Evidence for the Decoupling of Attention from Perceptual Input during Offline Thought. PLoS ONE. 6, e18298 (2011).
8. Walcher, S., Körner, C., Benedek, M.: Looking for ideas: Eye behavior during goal-directed internally focused cognition. Conscious. Cogn. 53, 165–175 (2017).
9. Sullivan, Y.: Costs and Benefits of Mind Wandering in a Technological Setting: Findings and Implications, https://digital.library.unt.edu/ark:/67531/metadc862836/.
10. Sullivan, Y., Davis, F., Koh, C.: Exploring Mind Wandering in a Technological Setting. ICIS 2015 Proc. (2015).
11. Bixler, R., D'Mello, S.: Toward Fully Automated Person-Independent Detection of Mind Wandering. In: Dimitrova, V., Kuflik, T., Chin, D., Ricci, F., Dolog, P., and Houben, G.-J. (eds.) User Modeling, Adaptation, and Personalization. pp. 37–48. Springer International Publishing (2014).

12. Zhao, Y., Lofi, C., Hauff, C.: Scalable Mind-Wandering Detection for MOOCs: A Webcam-Based Approach. In: Data Driven Approaches in Digital Education. pp. 330–344. Springer, Cham (2017).
13. Bixler, R., D'Mello, S.: Automatic Gaze-Based Detection of Mind Wandering with Metacognitive Awareness. In: Ricci, F., Bontcheva, K., Conlan, O., and Lawless, S. (eds.) User Modeling, Adaptation and Personalization. pp. 31–43. Springer International Publishing (2015).
14. Borlund, P.: The concept of relevance in IR. J. Am. Soc. Inf. Sci. Technol. 54, 913–925 (2003).
15. Freund, L., Kopak, R., O'Brien, H.: The effects of textual environment on reading comprehension: Implications for searching as learning. J. Inf. Sci. 42, 79–93 (2016).
16. Smallwood, J., Schooler, J.W.: The restless mind. Psychol. Bull. 132, 946–958 (2006).
17. Smallwood, J., Schooler, J.W., Turk, D.J., Cunningham, S.J., Burns, P., Macrae, C.N.: Self-reflection and the temporal focus of the wandering mind. Conscious. Cogn. 20, 1120–1126 (2011).
18. Feng, S., D'Mello, S., Graesser, A.C.: Mind wandering while reading easy and difficult texts. Psychon. Bull. Rev. 20, 586–592 (2013).
19. Witten, I.H., Frank, E., Hall, M.A., Pal, C.J.: Data Mining, Fourth Edition: Practical Machine Learning Tools and Techniques. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (2016).
20. Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.: SMOTE: Synthetic Minority Over-sampling Technique. ArXiv11061813 Cs. (2011).
21. Breiman, L.: Random Forests. Mach. Learn. 45, 5–32 (2001).