Automatic Detection of Mind Wandering from Multimodal Datastreams: A Survey of State-of-the-art Methods

Murtadha Al Nahadi, Robin Faber, Justin de Haan, Wessel Turk
Delft University of Technology
Delft, Netherlands
{M.AlNahadi, R.J.Faber, J.C.deHaan, W.R.A.Turk}@student.tudelft.nl

***Abstract*—This study aims to give a clear and structured overview of the current state-of-the-art of automatic detection of mind wandering from multimodal datastreams. This survey is a literature study done by finding and analyzing existing studies regarding the subject. From an initial 72 studies, a total of 18 were selected to be analyzed in our study. A technical overview of all evaluated studies is given in Table I. This table, alongside the context of the studies, is then further analyzed in the corresponding text. Considering how new the automatic detection of mind wandering is as a research subject, there have been a lot of studies using different detection technologies with good success. This shows that there are good opportunities for real-world applications of these technologies that can help prevent mind wandering. However, it would be beneficial for future research to have standardized reporting methods and datasets such that performance comparisons are easier to make.**

***Index Terms*—Mind wandering, automatic detection, attentional state estimation, machine learning**

I. Introduction

Mind wandering (MW) is a phenomenon where the thoughts of a person shift from task-related to task-unrelated for a certain period of time. Every person has probably experienced it at some point and it can come paired with inefficiency and distraction. In situations where attention is needed, such as driving, it is important that the mind wanders as little as possible so people can focus on the task at hand. Recent research has clearly shown that inattention when driving has an indisputable impact on road safety [1]. It is clear that automated detection and prevention of mind wandering in such settings could be a good solution to the problem.

This study aims to identify possible challenges in the field of automatic detection of MW in order to further develop it or to be used as a starting point for people who wish to do research and want an overview of the possibilities and what has been done already. This goal is accomplished by giving a structured overview of the state-of-the-art methods in the field and reasoning about the broader context of these studies.

We investigated the following aspects of existing research: In which tasks is the automatic detection of MW being utilized? Which modalities are used? Which sensors are used for recording these modalities and which features are extracted from them? How is MW being reported by the participants? Which machine learning (ML) algorithms are used to train the model and what performance is achieved by the model? In which environment has the data been collected and has this data been made publicly available?

II. Method

In order to identify all studies that are relevant, we used a systematic approach that can be broken down into six steps (Fig. 1).

First, 10 studies were hand-picked that were identified as relevant for our goal. These hand-picked studies served as a baseline for the upcoming search of online libraries. A systematic search on Scopus and Web of Science was performed to identify all studies aimed at the detection of mind wandering. The retrieved studies were checked on relevancy and completeness. An indication of

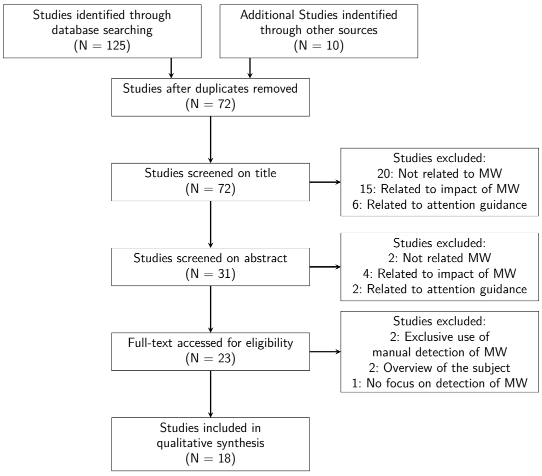

Figure 1: Flow Diagram outlining the screening process

the completeness of the search terms used was obtained by comparing the results of the online libraries with the baseline. The used search terms were tweaked by making them broader when the results contained a low number of studies and few baseline studies, and made narrower when there were many irrelevant studies retrieved. A logbook of the used search terms can be found in Appendix A. As a result of this first step, an initial list of 125 studies was obtained.

Second, all duplicates obtained from the different libraries were removed, leaving 72 studies.

Third, from this point onward, all remaining studies were divided into two stacks. Each stack was reviewed individually by two team members. The studies were screened on their title and were excluded if team members reached consensus on their irrelevancy. During the filtering process, we used the following three exclusion criteria:

1) **Not related to Mind Wandering**, e.g. [2], focuses on different formats to present lecture material and is therefore excluded.
2) **Related to the impact of Mind Wandering**, e.g. [3], concludes that mind wandering while driving leads to dangerous behaviour behind the wheel and is thus excluded.
3) **Related to attention guidance**, e.g. [4], explores gamification to restore students' attention level in classroom teaching and is thus excluded.

In case of uncertainty, studies were included to be screened on abstract. This step narrowed the list of studies down to 31.

Fourth, Eight studies were excluded while screening on abstract, leaving 23 studies remaining. The excluded studies fit the earlier defined exclusion criteria but looked interesting based on their title.

Fifth, every remaining article was accessed and screened on their full-text. This included reading the methodology, results, conclusion and inspecting their figures and tables. Two papers were excluded because they give an overview of the field, rather than explaining (new) methods of detecting mind wandering. Another two studies were excluded because they use manual reporting exclusively.

At last, one article was excluded because it describes a system that uses third-party solutions to detect mind wandering. The focus of this article lays on implementing this system in a learning environment, rather than on combining/improving detection methods for mind wandering. In conclusion, a total of 18 studies are left to review.

Finally, the studies were carefully studied, extracting relevant data. This data is then further analyzed.

III. Results & Discussion

A total of 18 studies were included in our review, a summary of the studies is shown in Table I. For each study, the following was extracted: the task people had to perform, the modalities used, the extracted features, used sensors, how people reported MW, the machine learning (ML) algorithms that were used and the performance of the developed classifier. In the following sections, the contents and relevance of the columns from Table I will be discussed.

Table I: Detection methods for mind wandering

Reference	Task	Modalities used	Sensors Used	Features Extracted [1]	Reporting MW [2]	ML algorithm(s)	Performance	Notes
Bixler et al. (2015) [5]	Reading	Eye gaze, physiology, context	Eye tracker, wrist sensor for physiology	GGF (46), LGF (23), SC (43), ST (43), CF (11)	WP, EP (both auditory)	13 supervised ML classifiers	Kappa of 0.19	Larger gaze window size results in a higher Accuracy
Bixler et al. (2015) [6]	Reading	Eye gaze	Eye tracker	GGF (46), LGF (20)	Pseudo-random AP, SCR	10 ML classifiers	Kappa of 0.45, Accuracy of 74%	Context features did not improve classification, SVM classifier works best, larger window sizes improved accuracy
Bixler et al. (2014) [7]	Reading	Eye gaze, context	Eye tracker	GGF (30), LGF (19), CF(11)	WP, EP (both auditory)	20 supervised ML classifiers	Kappa of 0.28, Accuracy of 72% (end-of-page), Kappa of 0.17, Accuracy of 59% (within-page)	Bayes net and naïve bayes classifiers worked best
Bixler et al. (2016) [8]	Reading	Eye gaze, context	Eye tracker	GGF (46), LGF (23), CF (11)	WP, EP (both auditory)	20 supervised ML classifiers	Kappa of 0.31, Accuracy of 72% (end-of-page), Kappa of 0.18, Accuracy of 67% (within-page)	Bayes net and naïve bayes classifiers worked best, window size had no effect
Blanchard et al. (2014) [9]	Reading	Physiology, context	Wrist sensor for physiology	SC (43), ST (43), CF (11)	WP, EP (both auditory)	Unspecified amount of supervised classifiers	Kappa of 0.22 (within-page), Kappa of 0.14 (end-of-page)	LADTree classifier performed best
Gwizdka (2019) [10]	Reading	Eye gaze	Eye tracker	27 eye gaze features	Pseudo-random WP (visual)	ML classifiers	Accuracy of 89%	Random forest classifier performed best, Accuracy might not be reliable due to resampling, window size had very little effect
Jo et al. (2017) [11]	Watching a video	Words in video	N/A	High frequency words	Not watching is compared with watching	N/A	N/A	Provides initial evidence that MW can be detected using high frequency words in video lectures
Stewart et al. (2016) [12]	Watching a video	Facial features and body movement	Webcam	Facial features (75), body movement features (3)	SCR	Several ML classifiers	F_1 score of 0.30	SVM classifier works best
Stewart et al. (2017) [13]	Watching a video	Facial features and body movement	Webcam	Facial features (75), body movement features (3)	SCR	9 supervised classification techniques	F_1 score of 0.39	Larger window sizes performed slightly better, SVM classifier works best

Continued on next page

Table I concluded from previous page

Reference	Task	Modalities used	Sensors Used	Features Extracted [1]	Reporting MW [2]	ML algorithm(s)	Performance	Notes
Zhao et al. (2017) [14]	Watching a video	Eye gaze	Webcam	GGF and LGF (58 in total)	SCR	Logistic Regression, Linear SVM, Naive Bayes	F_1 score of 0.40	Naive Bayes net works best
Zhang et al. (2016) [15]	Watching a video	Eye gaze, pupillometry	Camera of mobile device	Eye movement, pupil index	Participants review images afterwards	Logistic regression, Linear SVM, HMM	Accuracy of 53%	Linear SVM classifier works best, largest window size did not result in highest accuracy
Pham et al. (2015) [16]	Watching a video	Fingertip transparency, context	Mobile phone cameras	Heart rate features (12) and CF (7)	AP	Supervised ML classifiers	Kappa of 0.22, Accuracy of 71%	KNN classifier (K = 5) performed best
Hutt et al. (2017) [17]	Interacting with learning technology	Eye gaze, context	Eye tracker	GGF (57), LGF (81), CF (8)	Pseudo-random AP	Bayesian Networks	F_1 score of 0.59	-
Cheetham et al. (2016) [18]	Eyes-closed FA meditation	Biosignals	Biosensors	Respiration, heart rate, electrocardiogram, electromyogram, electrodermal acitivity	SCR	N/A	Accuracy of 85%	Performance achieved for area under the receiver operator characteristic curve
Da Silva et al. (2018) [19]	Maintain access to memory items (letters) while completing math equations	Mouse movements	Computer mouse	Time to press start, initiation time, total distance, average speed, speed errors	Probes after each set	N/A	N/A	Provides initial evidence that mouse tracking can be used to predict MW
Gontier (2016) [20]	Various extra-vehicular activities at the Mars Desert Research Station	Heart rate	ECG sensor	Measures of heart rate variability	AP	N/A	N/A	This study was only done with 1 participant, so did not mention much in terms of relevant results
Mishchenko et al. (2015) [21]	Controlling a passenger train in a train simulator program	Electroencephalographic brain activity	Modified EPOC EEG device	Feature vector containing 252 EEG spectral power values	AP	Several SVM classifiers	Accuracy of 90%	Accuracy is cross-validation accuracy
Russell et al. (2016) [22]	Watching flickers on a computer screen	Electroencephalographic brain activity	Electroencephalogram	ssVEP powers	N/A	N/A	N/A	An indication that high frequency ssVEPs are attention sensitive is given, due to unexpected results further research is needed

[1] *Features extracted:* GGF = global gaze features, LGF = local gaze features, SC = skin conductance, ST = skin temperature, CF = context features
[2] *Reporting MW:* WP = Within-page probes, EP = End-of-page probes, AP = Auditory probes, SCR = self-caught reports

A. Performed tasks

Table I shows a variety of different tasks that the participants had to perform, each with different advantages related to that study. The two most used tasks were reading a text and watching some type of content on a computer screen. These two methods are probably the most popular because they are easy to distribute to a large audience and do not require any advanced equipment. Other than these two tasks, none of them occurred more than once, because the others were usually specific to a certain study, such as interacting with learning software [17]. An interesting thing to note here is that the type of performed tasks did not seem to correlate with the modalities that were used. This can be seen in Table I, the task of watching something on a computer screen was evaluated in studies that recorded the modalities of eye gaze [14], EEG brain activity [22], facial features [13] and heart rate [16]. This suggests that MW can be automatically detected in a wide variety of ways within the same context, thus making the technology very flexible in terms of implementation.

B. Modalities recorded and features extracted

A modality is a channel of information provided by a sensor. For example, the modality physiology is used in [9]. This means that the study used some features of physiology to detect mind wandering. In this case, the study looked at the skin conductance and the skin temperature of the test subject. There are a lot of different modalities used in the studies that were looked at. A few modalities that were used are words in videos [11], pupillometry [15] and EEG brain activity [21][22]. The most common modality used in these studies is eye gaze. This could be because eye gaze has already shown promising results in various studies. It is also one of the easiest things to track in terms of equipment, seeing as it only requires a simple webcam, unlike modalities such as brain activity, which require expensive EEG scans.

To be able to do the research, the researchers have to extract features from the modalities. In Table I, the features extracted from modalities are listed. The abbreviations of the features can be found underneath the table. Every study was precise in describing the features that were extracted. Specifying exactly which features are extracted gives a better overview of what features might be important for the detection of MW. In one of the studies it is mentioned that 57 global gaze features were extracted [17]. Global gaze features are independent of what people are looking at, while local gaze features do take this into account. It is not clear which features result in better performance. In one study a combined model performed best and global gaze features outperformed local gaze features [7]. Another study showed that there was almost no difference in performance for using global gaze features or local gaze features and the combination of both performed worse [17]. Features that are extracted the most when looking at physiology are skin conductance and skin temperature. Other features that are often used are context features. These do not belong to a modality that is recorded by a sensor. These features are related to the task the participant has to perform (e.g., page length or task difficulty).

C. Sensors used

The same modalities can sometimes be captured using different sensors. Eye gaze, for example, is often measured using an eye tracker. Eye trackers are often expensive. If MW can only be automatically detected using expensive eye trackers, less people will use systems that can automatically detect MW. Expensive eye trackers are not always used however, in one study a commercial off the shelf eye tracker is used [17]. In order to make the automatic detection of MW more widely available it was also studied whether this could be done using a low-end webcam [12][13][14], or using the camera of a mobile device [15]. Another example is the heart rate, heart rate is measured using an ECG sensor [20], while this is also done using the cameras of a mobile phone [16]. There are also modalities that were always recorded using the same type of device. For example, physiology was always measured using a wrist sensor [5][9]. The fact that it is being studied whether different types of devices can detect MW shows that the researchers are aware of the impact of the types of sensors used on the applications of the detection of MW.

D. Reporting MW

The biggest problem with detecting mind wandering is that it is an internal state of the mind and is thus something that can not easily be measured. This makes the process of collecting training data difficult because it relies on alternate methods to gather information about whether a person is mind wandering.

In most papers, this was done with some type of probe that was activated while the test subject was performing a particular task. A probe in this context means a signal to the test subject, who then records whether he/she was mind wandering. This is a simple binary yes/no answer, no intensity of the MW is given in any of the studies. It is stated that probes are the most standard way for reporting MW because alternatives like EEG and fMRI have not been validated and are often not practical [5]. The most common type of probes were within-page and end-of-page probes (only applicable when the task was reading) and auditory probes, which give a certain sound when the test subject has to report whether or not they were mind wandering.

Another technique for reporting MW that was commonly used is self-caught reports (SCR). With this technique, the test subject has to report the MW the moment he/she becomes aware of it, which means they report a form of MW that occurs with metacognitive awareness [6]. An obvious downside to SCR is that there is no information about the time where no reports of MW occur. The test subject could be paying attention, but could also be mind wandering without realizing it yet [6]. This is a problem that does not occur when using probes, since the signal indicates when the subject reports, thus always guaranteeing a yes or no answer. Although this is a clear limitation of SCR, both the probe-caught and self-caught methods have been validated in a number of studies [5].

In terms of performance, both the different types of probes and SCR showed success in different studies, making either of them viable options to construct training data.

Another thing that is important when gathering data about MW is how much time is recorded of the MW phase, this is called the window size. Most studies mentioned the window sizes used, there were a lot of varying window sizes, anywhere from 1 to 60 seconds. Some studies also tried the same experiment with various window sizes. A study ran their experiment with window sizes 2, 8 and 16 seconds, a window size of 8 seconds gave the highest accuracy [15]. Other studies found higher accuracies for larger window sizes [5][6][13]. Overall, the optimal window size is probably very dependent on the study, so a conclusion about the best general window size to detect mind wandering cannot be drawn.

E. Performance of various ML algorithms

Different ML algorithms were used to train models that can detect MW. Many ML algorithms are used for the classification tasks because it is not yet known which algorithm is best suited to train a model that can automatically detect MW. Often, the ML algorithms that are used are from Weka[1] [5][6][8][7][9][10][17][16], a collection of machine learning algorithms. From these algorithms the best one(s) would be picked. Most studies concluded that Support Vector Machines and Bayesian models performed best. There were also studies that did not use ML algorithms, initial evidence was provided that high-frequency words in lecture videos could be used to detect MW [11]. These studies investigated whether there is a relation between specific features and MW. It also occurred that ML could not be applied due to the fact that there was too few participants in the study [20]. This resulted in a dataset that was too small to train a model with.

The performance was measured using different metrics of accuracy, for example by using a F_1 score, Cohen's kappa or as a percentage based accuracy. Percentage based accuracy can be defined in different ways, which is mostly dependent on the application. For example, percentage based accuracy was defined in one study as the area under the receiver operator characteristic curve [18], but cross-validation accuracy was also used [21]. This makes it hard to compare different studies. Some properties were however also compared within the study. Some results show that there is no significant difference in accuracy between a window size of 5 and 10 seconds [10]. This

[1]https://www.cs.waikato.ac.nz/ml/weka/

confirms that the optimal window size is not the same for every application and study. This study also showed a significant increase in accuracy over what was achieved in [7]. As mentioned before, it is important to verify whether MW could also be detected by using equipment that is owned by more people. One study found that a low-end webcam performed as good as an eye tracker [14]. Yet another study achieved an accuracy of 53% using a mobile device [15], showing a slight decrease in accuracy compared to similar studies using an eye tracker.

F. Published datasets

While reviewing the literature, it was questioned whether there are any published datasets for research concerning the automatic detection of MW. It was chosen to leave this column out of the table since only one of the studies mentioned a published dataset or published their own dataset [14]. Publishing datasets could be beneficial for the research in general. By publishing data other people could research different approaches of detecting MW without having to collect the data themselves. The data could also be used to validate the research that already has been done. Publishing datasets would also make it easier to test whether a certain ML algorithm that has not been tried in the study itself would result in a better model.

G. Environment

All studies were performed in some kind of experimental setting where participants are unfamiliar with their surroundings and/or equipment. Personal experience and the current environment determines how much *interference effects* a distractor will generate [23]. This interference effect explains why it is not always possible to keep your attention on a target. Since participants are in an unfamiliar environment and do not use their own equipment, they could experience *more* interference than normal. Contrariwise, because participants know their performance is being measured, it could also be that they are more focused, leading to *less* interference. The amount of interference could influence when, and even if a participant self-report their mind wandering. This has a big effect on the training of ML models and their results. It would be interesting to perform the studies in a more natural environment to see the effects this interference could have.

IV. Conclusion

While reviewing the literature, it was questioned whether there are any published datasets for research concerning the automatic detection of MW. It was chosen to leave this column out of the table since only one of the studies mentioned a published dataset or published their own dataset [14]. Publishing datasets could be beneficial for the research in general. By publishing datasets other people could research different approaches of detecting MW without having to collect the data themselves. The data could also be used to validate the research that already has been done. Publishing datasets would also make it easier to test whether a certain ML algorithm that has not been tried in the study itself would result in a better model.

References

[1] G. Berthié, C. Lemercier, P.-V. Paubel, M. Cour, A. Fort, C. Galéra, E. Lagarde, C. Gabaude, and B. Maury, "The restless mind while driving: drivers' thoughts behind the wheel," *Accident Analysis & Prevention*, vol. 76, pp. 159–165, 2015.

[2] K. E. Wilson, M. Martinez, C. Mills, S. D'Mello, D. Smilek, and E. F. Risko, "Instructor presence effect: Liking does not always lead to learning," *Computers & Education*, vol. 122, pp. 205–220, 7 2018.

[3] D. A. Albert, M. C. Ouimet, J. Jarret, M.-S. Cloutier, M. Paquette, N. Badeau, and T. G. Brown, "Linking mind wandering tendency to risky driving in young male drivers," *Accident Analysis and Prevention*, vol. 111, pp. 125–132, 2018.

[4] S. Xiao, W. Liang, and Y. Tang, "Classroom attention restoration using computer game rewarding mechanism," in *13th International Conference on Computer Science and Education, ICCSE 2018*, 2018, pp. 781–786.

[5] R. Bixler, N. Blanchard, L. Garrison, and S. D'mello, "Automatic detection of mind wandering during reading using gaze and physiology," in *ICMI 2015 - Proceedings of the 2015 ACM International Conference on Multimodal Interaction*, 2015, pp. 299–306.

[6] R. Bixler and S. D'Mello, "Automatic gaze-based detection of mind wandering with metacognitive awareness," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9146, pp. 31–43, 2015.

[7] R. Bixler and S. D'Mello, "Toward fully automated person-independent detection of mind wandering," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8538, pp. 37–48, 2014.

[8] R. Bixler and S. D'Mello, "Automatic gaze-based user-independent detection of mind wandering during computerized reading," *User Modeling and User-Adapted Interaction*, vol. 26, no. 1, pp. 33–68, 2016.

[9] N. Blanchard, R. Bixler, T. Joyce, and S. D'Mello, "Automated physiological-based detection of mind wandering during learning," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8474 LNCS, pp. 55–60, 2014.

[10] J. Gwizdka, "Exploring eye-tracking data for detection of mind-wandering on web tasks," in *Lecture Notes in Information Systems and Organisation*, vol. 29, 2019, pp. 47–55.

[11] J. Jo and H. Lim, "A Study of Keywords Based on the Word Frequency Effect Theory in Video Lectures of Software Engineering Education for Detecting Mind," in *Proceedings - 30th IEEE Conference on Software Engineering Education and Training, CSEE and T 2017*, vol. 2017-Janua, 2017, pp. 241–245.

[12] A. Stewart, H. Chen, N. Bosch, P. J. Donnelly, and S. K. D'Mello, "Where's your mind at? Video-based mind wandering detection during film viewing," in *UMAP 2016 - Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*, 2016, pp. 295–296.

[13] A. Stewart, N. Bosch, H. Chen, P. Donnelly, and S. D'Mello, "Face forward: Detecting mind wandering from video during narrative film comprehension," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10331 LNAI, pp. 359–370, 2017.

[14] Y. Zhao, C. Lofi, and C. Hauff, "Scalable mind-wandering detection for MOOCs: A webcam-based approach," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10474 LNCS, pp. 330–344, 2017.

[15] C.-l. Zhang and R.-m. Shen, "Automatic Detection of Mind Wandering Using Mobile Device In the Scenario of Online and Mobile Education," in *COMPUTER SCIENCE AND TECHNOLOGY (CST2016)*, N. Cai, Ed., 2017, pp. 153–162.

[16] P. Pham and J. Wang, "Attentivelearner: Improving mobile MOOC learning via implicit heart rate tracking," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9112, pp. 367–376, 2015.

[17] S. Hutt, C. Mills, N. Bosch, K. Krasich, J. Brockmole, and S. D'mello, "Out of the Fr-"Eye"-ing Pan: Towards gaze-based models of attention during learning with technology in the classroom," in *UMAP 2017 - Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*, 2017, pp. 94–103.

[18] M. Cheetham, C. Cepeda, and H. Gamboa, "Automated detection of mind wandering: A mobile application," in *BIOSIGNALS 2016 - 9th International Conference on Bio-Inspired Systems and Signal Processing, Proceedings; Part of 9th International Joint Conference on Biomedical Engineering Systems and Technologies, BIOSTEC 2016*, 2016, pp. 198–205.

[19] M. R. D. Da Silva, M. Postma-Nilsenová, and F. Hermens, "Wandering mice, wandering minds: Using computer mouse tracking to predict mind wandering," in *CEUR Workshop Proceedings*, vol. 2265, 2018.

[20] C. Gontier, "How to prevent mind-wandering during en EVA? Presentation of a mind-wandering detection method using ECG technology in a Mars-analog environment," in *Proceedings of the International Astronautical Congress, IAC*, 2016.

[21] Y. Mishchenko and M. Kaya, "Detecting the attention state of an operator in continuous attention task using EEG-based brain-computer interface [EEG Tabanli Beyin Makine Arayüzü Kullanilarak Devamli Dikkat Gerektiren Görevde Operatörün Dikkat Durumunun Tespiti]," in *2015 23rd Signal Processing and Communications Applications Conference, SIU 2015 - Proceedings*, 2015, pp. 232–235.

[22] B. A. H. Russell, J. C. Russo, I. P. Warfield, and W. D. Casebeer, "Monitoring attention with embedded frequency markers for simulation environments," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9743, pp. 394–403, 2016.

[23] C. Roda and J. Thomas, "Attention aware systems: Theories, applications, and research agenda," *Computers in Human Behavior*, vol. 22, no. 4, pp. 557–587, 2006.

Appendix A

Database	Search String	Retrieved (Baseline)	Remarks
Scopus	TITLE("Attention aware systems: Theories, applications, and research agenda" OR "Exploring Eye-Tracking Data for Detection of Mind-Wandering on Web Tasks" OR "Face forward: Detecting mind wandering from video during narrative film comprehension" OR "Toward fully automated person-independent detection of mind wandering" OR "Scalable mind-wandering detection for MOOCs: A webcam-based approach" OR "Automatic gaze-based user-independent detection of mind wandering during computerized reading" OR "Automatic gaze-based detection of mind wandering with metacognitive awareness" OR "Automatic detection of mind wandering during reading using gaze and physiology" OR "Out of the Fr-Eye-ing Pan: Towards gaze-based models of attention during learning with technology in the classroom"	10 (10/10)	Baseline
Web of Science	TI=("Attention aware systems: Theories, applications, and research agenda" OR "Exploring Eye-Tracking Data for Detection of Mind-Wandering on Web Tasks" OR "Face forward: Detecting mind wandering from video during narrative film comprehension" OR "Toward fully automated person-independent detection of mind wandering" OR "Scalable mind-wandering detection for MOOCs: A webcam-based approach" OR "Automatic gaze-based user-independent detection of mind wandering during computerized reading" OR "Automatic gaze-based detection of mind wandering with metacognitive awareness" OR "Automatic detection of mind wandering during reading using gaze and physiology" OR "Out of the Fr-Eye-ing Pan: Towards gaze-based models of attention during learning with technology in the classroom" OR "Automated detection of mind wandering: A mobile application")	6 (6/10)	This count is lower because Web of Science is very strict in terms of acceptance, so some of the papers we identified as relevant are not on the site
Scopus	TITLE-ABS-KEY (("mind wandering" OR "zoning out" OR "attention loss" OR "Task unrelated thought") AND (("attentional state estimation" OR "State models" OR "attention tracking" OR "user modeling") OR ("attention aware interfaces" OR "attentive user interfaces" OR "smart environment")))	12 (5/10)	This query gave too few results due to the small amount of keywords, but most of them seem relevant, which indicates that the keywords used are good.
Web of Science	TS=("Mind wandering" OR "Zoning out" OR "Attention loss" OR "Task unrelated thought" OR "Mindless reading") AND TS=("Attentional state estimation" OR "State models" OR "Attention tracking" OR "Eye-tracking" OR "Pupillometry" OR "Gaze tracking" OR "Gaze registration" OR "User modeling" OR "user modeling" OR "Affect detection" OR "Attention" OR "Learning analytics" OR "Attention aware interfaces" OR "Attentive user interfaces" OR "Smart environment" OR "Attention aware systems" OR "Controlling Human Attention" OR "MOOCs" OR "Adaptive systems" OR "Notification systems" OR "Intelligent tutoring systems" OR "Attention-aware learning") and WC = ("COMPUTER SCIENCE CYBERNETICS" OR "COMPUTER SCIENCE INFORMATION SYSTEMS" OR "COMPUTER SCIENCE SOFTWARE ENGINEERING" OR "COMPUTER SCIENCE ARTIFICIAL INTELLIGENCE" OR "COMPUTER SCIENCE INTERDISCIPLINARY APPLICATIONS" OR "COMPUTER SCIENCE THEORY METHODS")	23 (5/10)	Considering that the database is somewhat stricter the result is reasonably good.
Scopus	TITLE-ABS-KEY (("Mind wandering" OR "Zoning out" OR "Attention loss" OR "Task unrelated thought" OR "Mindless reading") AND (("Attentional state estimation" OR "State models" OR (Attention AND Tracking) OR ((Automated OR Automatic) AND detection) OR "Eye-tracking" OR "Pupillometry" OR "Gaze tracking" OR Gaze OR "Gaze registration" OR "User modeling" OR "Affect detection" OR "Attention" OR "Learning analytics") OR ("Attention aware interfaces" OR "Attentive user interfaces" OR "Smart environment" OR "Attention aware systems" OR "Controlling Human Attention" OR "MOOCs" OR "Adaptive systems" OR "Notification systems" OR "Intelligent tutoring systems" OR "Attention-aware learning"))) AND (LIMIT-TO (SUBJAREA , "COMP") OR LIMIT-TO (SUBJAREA , "ENGI"))	74 (9/10)	This search query gave us a good amount of results, but we only recognised about half of those as relevant, most likely because of the large amount of keywords used with OR operators.

Table A1: Log of search queries used