Gaze-based Detection of Mind Wandering during Lecture Viewing

Stephen Hutt[1], Jessica Hardey[1], Robert Bixler[1], Angela Stewart[1], Evan Risko[2]
and Sidney K. D'Mello[1]
[1]University of Notre Dame, [2]University of Waterloo
118 Haggar Hall, Notre Dame, IN, 46556, USA
{shutt, sdmello}@nd.edu

ABSTRACT
We investigate the use of consumer-grade eye tracking to automatically detect Mind Wandering (MW) during learning from a recorded lecture, a key component of many Massive Open Online Courses (MOOCs). We considered two feature sets: stimulus-independent global gaze features (e.g., number of fixations, fixation duration), and stimulus-dependent local features. We trained Bayesian networks using the aforementioned features and students' self-reports of MW and validated them in a manner that generalized to new students. Our results indicated that models built with global features (F_1 MW = 0.47) outperformed those using local features (F_1 MW = 0.34) and a chance-level model (F_1 MW = 0.30). We discuss our results in the context of MOOC development as well as integrating MW detection into attention-aware MOOCs.

Keywords
eye-gaze, Massive Open Online Courses, lecture viewing, intelligent tutoring systems, mind wandering, attention-aware learning

1. INTRODUCTION
Imagine you are giving a lecture on population diversity, most of your audience is engaged; however, one or more of your students are displaying signs of inattentiveness (e.g., dozing off, staring blankly). You may call on such a student in the hope of bringing their attention back to the lecture. You may even suggest a short break if too many students appear to be inattentive. This adaptation to your lecture was only possible because you had the ability to continually monitor your students' levels of attentional focus and to alter your instruction in real-time.

Now imagine you are teaching a Massive Open Online Course (MOOC). Your students are no longer in the same room as you and in many cases are not viewing the lecture at the same time you are delivering it. You no longer have the ability to monitor students' attentional focus and adapt to signs of inattentiveness.

Despite the challenges for educators, MOOCs are an increasingly popular method amongst students for e-learning and distance learning [16]. They have also been popular in traditional learning environments as alternate ways for delivering material [27]. MOOCs are often distributed world-wide to a variety of students across platforms with no limitations on individual participation. While there are some advantages to MOOCs with respect to promoting access, little is known with regard to how they address individual learners' needs. MOOCs have long had issues with extremely high dropout rates [1, 37], far greater than those in 'traditional' classroom environments. Though there has been work tying students' experiences with MOOCs to the dropout rate [37], there has been little exploration as to individual user experiences and trends that lead to retention problems [1, 17].

As a step towards better understanding student engagement within MOOCS, we focus on one form of disengagement called mind wandering (MW). MW is defined as an attentional shift from task-related processing towards internal task-unrelated thoughts [31]. In the context of learning, both lab and field studies have consistently reported MW rates in the 20%-50% range [21, 26, 34]; work looking at specifically recorded lectures showed the MW rates to be 20-45% [26, 34]. Additionally, a recent meta-analysis revealed a negative correlation between MW and performance across a variety of tasks [23]. MW negatively impacts a learner's ability to attend to external events [30], to encode information into memory [29], and to comprehend learning materials [28, 30]. As a result, MW is generally found to have a negative impact on learning outcomes.

Attempts to assuage the cost of MW rely on knowing if MW has occurred. However, detecting MW is no easy task. Although MW is related to other forms of disengagement, such as boredom, behavioral disengagement, and off-task behaviors [2, 3, 36], it is inherently distinct because it involves internal thoughts rather than overt expressive behaviors. This raises two challenges. First, while other disengaged behaviors often involve detectable behavioral markers (e.g., yawns signaling boredom), mind wandering is an internal state that can appear similar to being on-task [31]. Second, the onset and duration of MW cannot be precisely measured because MW can occur outside of conscious awareness [32].

Despite these challenges, there has been some progress toward automatic detection of mind wandering (discussed as related works in Section 1.1). However, almost all of the current MW detectors focus on reading. In contrast, we consider MW detection while students view MOOC-like lectures, building and validating the first gaze-based MW detector during video lecture viewing. We focus on video lectures because they are a core component of many courses and are vital to MOOCs. As MOOCs and lecture capture systems become more popular, we envision a variety of challenges with respect to keeping students engaged when content delivery occurs outside of the classroom with the instructor not even present. In this work, we harness the use of a computer in content delivery to take a step towards an attention-aware MOOCs.

1.1 Related Work

In an early study attempting to detect MW in the context of learning [10], students were asked to read aloud a paragraph about biology, followed by either self-explaining or paraphrasing. Students self-reported how frequently they zoned out on a scale from 1 (all the time) to 7 (not at all). Reports were then grouped as either low (1-3 on the scale) or high (5-7 on the scale). Supervised machine learning methods were trained using acoustic-prosodic features to classify these instances, achieving an accuracy of 64%. However, it is unclear whether this detector could generalize to new students as the validation method did not ensure student-level independence across training and testing sets.

Researchers have also built MW detectors based on information readily available in log files collected during the reading (e.g., reading time, complexity of the text). For example, [19] attempted to classify whether students were MW while reading a screen of text using reading behaviors and textual features (e.g., text difficulty). They were able to classify MW at 21% greater than chance using a leave-one-subject out cross-validation method. Similarly, another study [11] also attempted to predict MW during reading using textual features such as word familiarity, difficulty, and reading time. However, rather than using supervised machine learning, they used a set of researcher-defined thresholds to ascertain if participants were “mindlessly reading” based on difficulty and reading time.

More recent studies have explored additional techniques to detect MW during self-paced computerized reading [5, 8, 11]. In these studies, MW was measured via thought probes that occurred on pseudo-random screens (i.e. screen of text similar to a page of text). Participants responded either “yes” or “no” based on whether they were MW at the time of the probe. Supervised classification models were trained to discriminate the two responses using physiological features (e.g., skin conductance, temperature) [8] or eye-gaze [5], achieving accuracies ranging from 18% to 23% above chance and validated in a manner that generalized to new students. Further, combining the two modalities led to an 11% improvement in detection accuracy above the best individual modality [4].

Beyond reading, Pham et al. [22] provide initial proof that MW detection is possible during lecture viewing. Students watched video lectures on a smart phone using a MOOC-like application and responded yes or no to thought probes during the lectures. They used student heart rate (extracted via photoplethysmography) to train classifiers to detect MW. They achieved a 22% greater than chance detection accuracy, thereby providing some initial evidence of MW detection in a MOOC-like learning environment.

Hutt et al. [15] focused on detecting MW during learning with an intelligent tutoring system (ITS). Students’ eye gaze was tracked with a consumer grade eye tracker as they completed a 30-40 minute learning session with the ITS. Students reported MW by responding to pseudo-random thought probes throughout the session. A variety of supervised classification models were trained to detect MW from eye movements and basic contextual information (e.g., time within session), achieving student-independent MW detection that was 37% greater than chance.

Finally, Mills et al. [18] studied MW detection in the context of viewing a narrative film. This study used a research grade eye tracker to monitor eye movements from which content-free global gaze features (e.g., fixation duration) as well as content specific features were computed. The content specific features were generated from two areas of interest (AOIs): one from the saliency map of the image [14], and one specific to the film being watched. These AOIs were then used in conjunction with eye gaze to generate content specific (local) features (e.g., average distance of fixations from an AOI or intersections with the AOI). The key finding was that, unlike in reading tasks, models built using local features were more successful than those built from global gaze features, achieving a student-independent score of 29% above chance.

1.2 Current Study and Novelty

The novelty of this paper is two-fold. First, we build the first gaze-based detector of MW during video lecture viewing. We focus on eye tracking due to well-known relationships between visual attention and eye-movements. For example, MW has been associated with longer fixation durations [25] and more blinking in reading [33]. We use low-cost consumer-grade eye trackers to collect gaze data from participants as they view a recorded lecture (see Figure 1). Since research grade eye trackers can cost upwards of $40,000, the selection of affordable equipment (less than $150) increases the applicability of this work, enabling its eventual deployment in real world learning environments such as classrooms or students’ homes.

Second, we compare MW detection with the more generalizable, global eye gaze features to AOI based local features. Global eye gaze features have previously been successful for detecting MW in learning contexts such as reading [7] and interacting with an ITS [15]; however, recent work involving narrative film comprehension found that AOI based features were more effective in that context [18]. We explore if the differences in visual style and production techniques between a recorded lecture (Figure 1) and a narrative film (Figure 2) influence the effectiveness of local features for detecting MW. This is a critical comparison because the global features are much more generalizable.

Figure 1. Example frame from recorded lecture

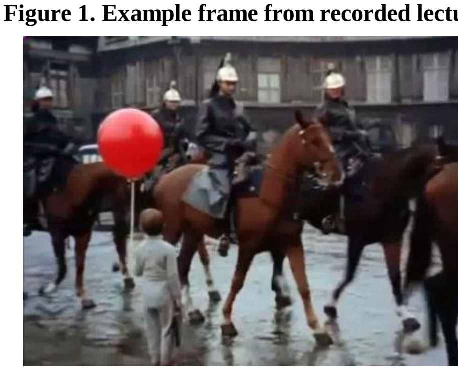

Figure 2. Example frame from narrative film

2. MW DETECTION

2.1 Procedure

Participants (or students) were 32 undergraduate students from a Canadian University, and they were compensated with course credit for their participation in the study. Participants watched a 24 minute lecture on population growth and were informed that there would be a test over what they had learned after watching the video. MW was defined as "Any thoughts that are not related to the material being presented", with examples such as "Concerns about an upcoming exam" and "Thoughts about dinner". Students also had the opportunity to ask questions regarding the instructions before the video began, but throughout the process, students had no control over the video.

Eye movements were monitored using a COTS eye-tracker called the EyeTribe that retails for $99. The eye tracker was placed just below the monitor on the desk.

2.2 Thought Probes

Mind wandering was measured during the recorded lecture using auditory thought probes, which is a standard approach in the literature [30]. Each student received 12 probes throughout the course of the recorded lecture that appeared at pre-determined times in the video. For each probe, the video paused and text was displayed on the screen asking, "In the moments prior to the probe were you MW?" Participants could then respond "1" for yes or "0" for no. Overall 31% of the probes were MW.

It is important to emphasize a few points about the method used to track MW. First, this method relies on self-reports because MW is an inherently internal phenomenon which requires self-awareness for reporting [32]. Second, self-reports of MW have been objectively linked to patterns in pupillometry [12], eye-gaze [25], and task performance [23], providing validity for this approach. However, at this time, there are no reliable neurophysiological or behavioral markers that can accurately substitute for the self-report methodology [32]. Indeed, this is the very reason we set out to build gaze-based MW detectors. The limits of thought probes are considered further in the Discussion section. For now, we note that our use of thought-probes to measure MW is consistent with the state of the art in the psychological and neuroscience literatures [32].

2.3 Feature Engineering

We calculated features from 30-second windows (window size was based on previous work [6, 15]) preceding each thought probe. We investigated two types of features: global gaze (from previous work [15]) as well as local features (based on [18]). Global gaze features focus on general gaze patterns and are independent of the content on the screen; whereas, local features encode where gaze is fixated on the screen.

2.3.1 Global Features

Eye movements were measured by fixations (i.e., points in which gaze was maintained on the same location) and saccades (i.e. the movement of the eyes between fixations). We calculated fixations and saccades from the raw eye gaze data using the Open Gaze and Mouse Analyzer (OGAMA) [35]. We considered six general measures across the 30-second window (bolded in Table 1) from which we computed the number, mean, median, minimum, maximum, standard deviation, range, kurtosis, and skew of the distributions, yielding 54 features. We also included three other features (see Table 1), yielding a total of 57 global gaze features.

Table 1. Eye-gaze features. Bolded cell indicates that nine descriptives (e.g., mean) were used as features (see Text)

Feature	Description
Fixation Duration	Elapsed time in ms of fixation
Saccade Duration	Elapsed time in ms of saccade
Saccade Length	Distance of saccade in pixels
Saccade Angle Absolute	Angle in degrees between the x-axis and the saccade
Saccade Angle Relative	Angle of the saccade relative to previous gaze point.
Saccade Velocity	Saccade Length / Saccade Duration
Fixation Dispersion	Root mean square of the distances of each fixation to the average fixation position
Horizontal Saccade Proportion	Proportion of saccades with relative angles <= 30 degrees above or below the horizontal axis
Fixation Saccade Ratio	ratio of fixation duration to saccade duration

2.3.2 Local Features

Local features were computed based on the relationship between eye movements and an area of interest (AOI). Two AOIs were defined for each frame of the lecture video that fell within the window: the most visually salient region of the frame, and the face of the lecturer. Visual saliency was determined using a MATLAB implementation of the Graph-Based Visual Saliency Algorithm [14] which produced a saliency map of pixel intensity from 0 to 1 for each frame that considered color, intensity, orientation, contrast, and movement. Determining the most visually salient region consisted of removing pixels with an intensity below a certain threshold (starting at 60% of the most intense pixel in the frame), leaving one or more regions of pixels as seen in Figure 4.

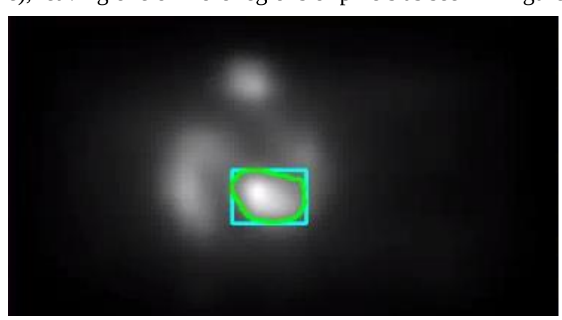

Figure 3. Example most salient region, lighter areas indicate higher saliency.

If the largest region had an area less than 2000 pixels (about 2% of the total area and a similar size to the face AOI), it was selected as the most visually salient region; otherwise, the process was repeated with a lower threshold. Figure 3 shows an example selection; in this case, the lecturer is gesturing, and the hand area was chosen as the most salient region. The face AOI was computed by detecting the facial location in the video using the commercially available software, Emotient [38]. The software provided the height and width of the face as well as the location which was converted into a bounding box after adding a small buffer of 20 pixels to account for any tracker inaccuracies.

There were 17 features calculated from each AOI for a total of 34 features. The features can be divided into three types: (1) AOI distance, (2) AOI intersection, and (3) saccade landing. AOI distance features consisted of descriptive statistics (minimum, maximum, mean, median, standard deviation, skew, kurtosis, and range) of the distance between the center of the AOI and the fixation position for each frame where the AOI was present, for a total of eight AOI distance features per AOI. AOI intersection features captured the proportion of time that gaze was within the bounding box or within one or two degrees of visual angle from the bounding box, resulting in a total of three AOI intersection features per AOI. Saccade landing features consisted of counting the number of times saccades landed on an AOI, left an AOI, or occurred within an AOI. To account for tracking noise, an additional set of saccade landing features were computed that counted the same events if they occurred within one degree of visual angle from the AOI, for a total of six saccade landing features per AOI.

2.4 Model Building

We focused on Bayesian Networks as they yielded the best performance compared to several other standard classifiers on this task in our previous work [15]. We used the default implementation from the Weka data mining package [13]. We validated the models with a leave-one-participant-out cross-validation scheme. For each fold, probe responses of one participant are held out for testing, and the model is trained on the remaining probes. This process ensures that no instances of any individual participant could appear in both the training and testing sets within a fold. This process is then repeated for the number of participants.

In total, there were 384 probes during the lecture. Of those, 12 were discarded due to insufficient eye gaze data (< 1 fixation) in the respective window to compute all the global features. The remaining 372 instances were used across all feature sets to ensure a fair comparison. Students reported MW in 31% of the 372 instances, thereby leading to data skew. This imbalance between labels poses a challenge as supervised learning methods tend to bias predications towards the majority class label. To compensate for this concern, we use the SMOTE algorithm [9] to create synthetic instances of the minority class by interpolating feature values between an instance and its randomly chosen nearest neighbors until the classes were equated. SMOTE was *only done on the training sets;* testing sets were unaltered in order to ensure validity of the results.

2.5 Results

The classification results are shown in Table 2. Because our intention is to detect instances of MW, we focus on the precision, recall, and F_1 score of the MW class as our key metric. For comparison, a chance-level baseline was created by *randomly* assigning the MW label to 31% (i.e., the MW baserate) of the instances over 1,000 iterations and averaging the result.

The results indicated that, while all models outperform the chance baseline: (1) global features outperformed local features and (2) adding local features to the global features increased precision but decreased recall, leading to no improvement in F_1 MW over global features alone. The fact that the best results were obtained from global features is significant because these features are more likely to generalize across interaction contexts.

Table 2. MW detection results for the recorded lecture

Feature Set	F_1 MW	Precision MW	Recall MW
Global	0.47	0.39	0.62
Local	0.36	0.40	0.34
Global + Local	0.42	0.45	0.39
Chance	*0.30*	*0.30*	*0.30*

3. GENERAL DISCUSSION

MOOCs present an exciting new era for education, providing more resources for traditional and non-traditional students alike. However, little is known about user experience and student engagement [17] with MOOCs, and it is widely known that they are plagued with poor retention rates [37]. Attention is critical to learning, [23] and monitoring attentional states of students is a step towards better understanding the learning process. MW is one key attentional state that is negatively correlated with learning [21]. MW is a covert, internal state with no obvious behavioral markers, making it difficult to detect. Although strides have been made to detect MW using eye gaze in the context of self-paced reading, gaze-based MW detection has not yet been attempted in the context of recorded lectures, a key component of many MOOCs. This is a challenge we address in the current paper. In the remainder of this section, we discuss our main findings, potential applications, and discuss limitations and future work.

3.1 Main Findings

MW detection during reading is supported by decades of research on attention and eye movements [24]. Recent work has branched away from reading into more complex environments [15, 18] that are not afforded with predictable patterns of eye moments. We have shown that MW detection is possible in the context of viewing a recorded lecture. We were able to accurately classify MW with an F_1 of 0.47 which is a 56% improvement over chance. Although this result is modest, it is an important first step in detecting MW in this domain, especially using consumer-grade eye tracking equipment.

Since MW detection in the context of online learning is still in its infancy, it is important that we explore techniques that are both successful and generalizable. We considered two feature sets in this work: global eye gaze features, which have previously performed well at detecting MW during reading and while interacting with an ITS, and local features, based on AOIs, that have previously been shown to be successful predicting MW during narrative film viewing. In the context of lecture viewing, we have shown that global eye movements outperform local AOI-based features, contrasting previous work during narrative film viewing [18] that found the opposite pattern.

It is interesting to consider why AOIs were less successful in this context as opposed to narrative film viewing. One suggestion lies in the different styles of the two media. Commercial, narrative films are directed with the viewer in mind, directing the audience's attention to whatever is pertinent. In many cases, films are produced by professionals with years of experience and numerous qualifications in their art form. In contrast, a recorded lecture involves far more basic film production techniques, and in many cases the film audience is the secondary audience; the lecture itself is designed for the audience in the room. Our methods rely on automated AOI detection. It may be that these style differences affect that detection, having a downstream effect on the features generated from those AOIs. Further research would be required to confirm this hypothesis.

All data was collected using low-cost, consumer-grade eye trackers (less than $150). This is a marked contrast compared to many research-grade trackers that can cost tens of thousands of dollars. Our hope is that these models can be deployed at scale and can be used to improve engagement and learning from MOOCs. For this reason, it was important to ensure that our models were validated in a student-independent manner which increases our models' ability to generalize to new students. The combination of student-independent models and consumer grade eye tracking increases our confidence that the models will generalize more broadly to applications outside of the laboratory, though this claim requires further empirical validation.

3.2 Applications

Lecture videos play a major role in online learning with MOOCs, so our MW detectors can be quite beneficial in that context. Our detectors could be implemented to provide real time updates to the MOOC software regarding the students' attention. Should a student be MW, the MOOC software could then adopt a variety of potential intervention strategies to refocus attention to the learning task. This could include simply pausing the video, asking a content-specific question, or asking the student to self-explain content that has recently been covered. Both interleaved questions [34] and self-explanations [20] have been shown to be effective in focusing attention. Students who answer incorrectly could then be encouraged to further review material and try again or could be redirected to an earlier point in the video. These approaches would give them multiple opportunities to correct the learning deficits attributed to MW.

It is important to consider that such interventions rely on MW detection which is inherently imperfect. The detector may issue a false alarm, suggesting that a student is MW when (s)he is not, or it could miss that a student is MW. In our view, MW detection does not need to be perfect as long as there is a modicum of accuracy. Imperfect detection can be addressed with a probabilistic approach, where the detector outputs a MW likelihood that is then used to determine whether an intervention is triggered (i.e., if the likelihood of MW is 70%, then there is a 70% chance of an intervention). The interventions should also be designed to "fail-soft" in that there are no harmful effects to learning if delivered incorrectly.

A further application is to inform the development of future MOOCs. Data from students' attention patterns whilst interacting with a MOOC video can be used to improve course structure (e.g. number of lectures and lecture length as well as course content such as individual explanations).

3.3 Limitations

We designed our approach to include a low-cost eye tracker, however, consumer models have a lower sampling-rate, limiting the accuracy of eye-gaze data compared to research-grade eye trackers. Furthermore, a key limitation was that we considered one lecture, so generalizability to other lectures is unknown. In addition, data was collected in a quiet lab environment; for better ecological validity we would need to explore more authentic learning environments (e.g. homes or libraries).

A further limitation relates to the use of thought probes which require users to be mindful of their MW and respond honestly. Although this methodology has been previously validated [12, 23, 25] there is no clear alternative to track a highly internal state like MW outside of measuring brain activity in an fMRI scanner. One futuristic possibility is to combine self-reports and wearable electroencephalography (EEG) as a means of collecting more accurate MW responses, but it is unclear if this can be done in more realistic contexts.

3.4 Future Work

The results discussed here invite several possibilities for improvement that we will address as future work. First, we will explore eye movements in different lectures. Having shown that global gaze models are applicable in this context, we will explore if we can train a model on one recorded lecture and use that model on other lectures and other topics. We will also explore cross training to other educational environments, to gain a better understanding of the differences and similarities in eye movements and attention across learning situations.

Another potential avenue is to integrate the detector into a MOOC to detect MW in real time. Here, the MW probes will be based upon the detectors real time assessment of students' attention instead of pre-prescribed or pseudo random probing. We can then better evaluate our detectors by comparing the probabilistic assessment of MW to students' responses to probes. Providing this refinement is successful, we could then use the detector to create a MOOC environment that intervenes in real time.

4. CONCLUSION

The popularity of MOOCs has ushered in an exciting time for students everywhere while also bringing challenges for educators. Advances in consumer grade eye tracking allow us to take a step towards a better understanding of how students engage with MOOCs on a larger scale. We have shown that we can detect MW in recorded lectures at above chance level. While much MW research has focused on the context of reading, our findings suggest that it might be possible to apply research on eye gaze, attention, and learning to this new context, thereby affording new discoveries about how students learn and interact with MOOCs while designing interfaces to sustain attention during learning.

5. ACKNOWLEDGMENTS

This research was supported by the National Science Foundation (NSF) (DRL 1235958 and IIS 1523091). Any opinions, findings and conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the NSF.

6. REFERENCES

[1] Adamopoulos, P. 2013. What Makes a Great MOOC? An Interdisciplinary Analysis of Student Retention in Online Courses. *International Conference on Information Systems* (2013), 21.

[2] Arroyo, I. et al. 2007. Repairing disengagement with non-invasive interventions. *Artificial Intelligence in Education* (Amsterdam, The Netherlands, 2007), 195–202.

[3] Baker, R.S.J. d. 2007. Modeling and understanding students' off-task behavior in intelligent tutoring systems. *SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2007), 1059–1068.

[4] Bixler, R. et al. 2015. Automatic detection of mind wandering during reading using gaze and physiology. *International Conference on Multimodal Interaction* (2015), 299–306.

[5] Bixler, R. and D'Mello, S. 2015. Automatic gaze-based user-independent detection of mind wandering during

computerized reading. *User Modeling and User-Adapted Interaction*. (2015), 1–36.

[6] Bixler, R. and D'Mello, S. 2016. Automatic gaze-based user-independent detection of mind wandering during computerized reading. *User Modeling and User-Adapted Interaction*. 26, 1 (2016), 33–68.

[7] Bixler, R. and D'Mello, S.K. 2014. Toward fully automated person-independent detection of mind wandering. *User Modeling, Adaptation, and Personalization* (Aalborg, Denmark, 2014), 37–48.

[8] Blanchard, N. et al. 2014. Automated physiological-based detection of mind wandering during learning. *Intelligent Tutoring Systems* (Switzerland, 2014), 55–60.

[9] Chawla, N.V. et al. 2002. SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*. 16, 1 (Jun. 2002), 321–357.

[10] Drummond, J. and Litman, D. 2010. In the zone: Towards detecting student zoning out using supervised machine learning. *Intelligent Tutoring Systems* (Pittsburgh, PA, USA, 2010), 306–308.

[11] Franklin, M.S. et al. 2011. Catching the mind in flight: using behavioral indices to detect mindless reading in real time. *Psychonomic Bulletin & Review*. 18, 5 (Oct. 2011), 992–997.

[12] Franklin, M.S. et al. 2013. Window to the wandering mind: pupillometry of spontaneous thought while reading. *The Quarterly Journal of Experimental Psychology*. 66, 12 (2013), 2289–2294.

[13] Hall, M. et al. 2009. The WEKA data mining software: An update. *SIGKDD Explorations*. 11, 1 (Nov. 2009), 10–18.

[14] Harel, J. et al. 2006. Graph-based visual saliency. *NIPS* (2006), 5.

[15] Hutt, S. et al. 2016. The eyes have it: gaze-based detection of mind wandering during learning with an intelligent tutoring system. *The 9th International Conference on Educational Data Mining* (Raleigh, NC, USA, 2016), 86–93.

[16] Liyanagunawardena, T. et al. 2013. MOOCs: A systematic study of the published literature 2008-2012. *The International Review of Research in Open and Distributed Learning*. 14, 3 (2013), 202–227.

[17] Milligan, C. et al. 2013. Patterns of engagement in massive open online courses. *Journal of Online Learning with Technology*. 9, 2 (2013), 149–159.

[18] Mills, C. et al. 2016. Automatic gaze-based detection of mind wandering during film viewing. *The 9th International Conference on Educational Data Mining.* (Raleigh, NC, USA, 2016).

[19] Mills, C. et al. 2015. Toward a real-time (day) dreamcatcher: sensor-free detection of mind wandering during online reading. *The 8th International Conference of Educational Data Mining* (Madrid, Spain, 2015), 786–789.

[20] Moss, J. et al. 2013. The nature of mind wandering during reading varies with the cognitive control demands of the reading strategy. *Brain Research*. 1539, (2013), 48–60.

[21] Olney, A.M. et al. 2015. Attention in educational contexts: The role of the learning task in guiding attention. *The Handbook of Attention*. J. Fawcett et al., eds. MIT Press.

[22] Pham, P. and Wang, J. 2015. AttentiveLearner: improving mobile MOOC learning via implicit heart rate tracking. *Artificial Intelligence in Education* (Madrid, Spain, 2015), 367–376.

[23] Randall, J.G. et al. 2014. Mind-wandering, cognition, and performance: a theory-driven meta-analysis of attention regulation. *Psychological Bulletin*. 140, 6 (Nov. 2014), 1411–1431.

[24] Rayner, K. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*. 124, 3 (Nov. 1998), 372–422.

[25] Reichle, E.D. et al. 2010. Eye movements during mindless reading. *Psychol Sci*. 21, 9 (Sep. 2010), 1300–1310.

[26] Risko, E.F. et al. 2013. Everyday attention: Mind wandering and computer use during lectures. *Computers & Education*. 68, (2013), 275–283.

[27] Sandeen, C. 2013. Integrating MOOCS into traditional higher education: The emerging "MOOC 3.0" era. *Change: The magazine of higher learning*. 45, 6 (Nov. 2013), 34–39.

[28] Schooler, J.W. et al. 2004. Zoning out while reading: Evidence for dissociations between experience and metaconsciousness. *Thinking and seeing: Visual metacognition in adults and children*. MIT Press. 203–226.

[29] Seibert, P.S. and Ellis, H.C. 1991. Irrelevant thoughts, emotional mood states, and cognitive task performance. *Memory & Cognition*. 19, 5 (Sep. 1991), 507–513.

[30] Smallwood, J. et al. 2008. When attention matters: the curious incident of the wandering mind. *Memory & Cognition*. 36, 6 (Sep. 2008), 1144–1150.

[31] Smallwood, J. and Schooler, J.W. 2006. The restless mind. *Psychological Bulletin*. 132, 6 (Nov. 2006), 946–958.

[32] Smallwood, J. and Schooler, J.W. 2015. The science of mind wandering: Empirically navigating the stream of consciousness. *Annual Review of Psychology*. 66, (2015), 487–518.

[33] Smilek, D. et al. 2010. Out of mind, out of sight: eye blinking as indicator and embodiment of mind wandering. *Psychological science*. 21, 6 (Jun. 2010), 786–789.

[34] Szpunar, K.K. et al. 2013. Mind wandering and education: from the classroom to online learning. *Frontiers in Psychology*. 4, (2013), 495.

[35] Vosskuhler, A. et al. 2008. OGAMA (Open Gaze and Mouse Analyzer): open-source software designed to analyze eye and mouse movements in slideshow study designs. *Behavior Research Methods*. 40, 4 (Nov. 2008), 1150–1162.

[36] Wixon, M. et al. 2012. WTF? detecting students who are conducting inquiry without thinking fastidiously. *User Modeling, Adaptation, and Personalization*. Springer. 286–296.

[37] Zheng, S. et al. 2015. Understanding student motivation, behaviors and perceptions in MOOCs. *Computer Supported Cooperative Work; Social Computing* (New York, NY, USA, 2015), 1882–1895.

[38] 2016. *Emotient module: Facial expression emotion analysis*.