Understanding and Detecting Divided Attention in Mobile MOOC Learning

Xiang Xiao[1, 2], Jingtao Wang[1, 2]
[1]Department of Computer Science, [2]Learning, Research and Development Center (LRDC)
University of Pittsburgh, Pittsburgh, PA, USA
{xiangxiao, jingtaow}@cs.pitt.edu

ABSTRACT
The emergence of mobile apps for Massive Open Online Courses (MOOCs) allows learners to access quality learning materials at low cost and "*to control where, what, how and with whom they learn*". Unfortunately, when compared with traditional classroom education, learners face more distractions and are more likely to multitask when they study alone in an informal learning environment. In this paper, we investigate the impact of divided attention (DA) on both the learning process and learning outcomes in the context of mobile MOOC learning. We propose OneMind, a system and algorithm for detecting divided attention on unmodified mobile phones via implicit, camera-based heart rate tracking. In an 18-participant study, we found that internal divided attention has a significant negative impact on learning outcomes; and that the photoplethysmography (PPG) waveforms implicitly captured by OneMind can be used to detect the presence, type, and intensity of divided attention in mobile MOOC learning.

Author Keywords
MOOC; Mobile interfaces; PPG; Divided attention.

ACM Classification Keywords
H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces.

INTRODUCTION
Massive Open Online Courses (MOOCs), as a scalable paradigm in online education, are experiencing rapid growth since 2011. In addition to the increased scale and reduced cost in knowledge dissemination, researchers found that MOOCs could help students develop new skills such as autonomous learning, and improve the visibility of universities and courses [9]. At the same time, challenges such as high dropout rates, and lack of individualized assessment and feedback also make today's MOOCs an inferior choice to one-on-one tutoring [3].

CHI 2017, May 06-11, 2017, Denver, CO, USA
© 2017 ACM. ISBN 978-1-4503-4655-9/17/05…$15.00
DOI: http://dx.doi.org/10.1145/3025453.3025552

Many MOOC providers such as Coursera, Udacity, and edX, have offered mobile apps to support "learning on the go". Unfortunately, mobile MOOC learning faces significant challenges when compared with classroom education or even traditional MOOCs on PCs. This is in part due to the highly diversified learning environments and highly interruptive learning contexts when learning with one's mobile device. Distractions could come from both *external* sources (e.g. background conversations or ambient noise) and *multitasking* (e.g. checking/updating social networking sites). When learners divide their attention between learning materials and other tasks or external distractions, the interference hampers their intentional use of memory [7] and reduces the memory performance substantially [4]. Both outcomes hinder the knowledge encoding process and lead to decreased understanding of the learning materials.

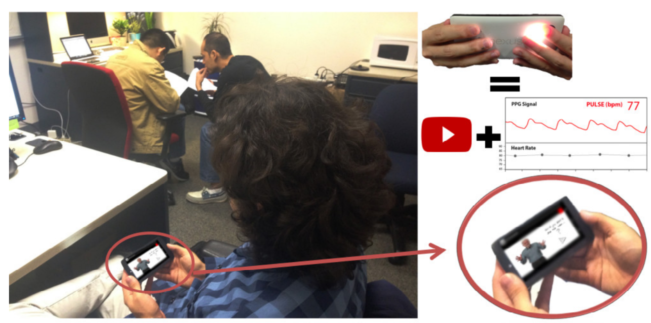

Figure 1. Using OneMind for MOOC learning while colleagues discuss in the background.

In this paper, we investigate the impact of divided attention (DA) on the learning process and learning outcomes in the context of mobile MOOC learning. We also propose OneMind (Figure 1), a mobile MOOC learning system that detects the presence, type, and intensity of divided attention via implicit physiological signal sensing on unmodified mobile phones. OneMind is built on top of AttentiveLearner [12, 18], monitoring a learner's heart rate implicitly via commodity camera based photoplethysmography (PPG) sensing. We present the design, implementation, and evaluation of OneMind. Through an 18-subject user study, we show the feasibility of using heart rate signals, implicitly recorded by the built-in camera of mobile phones, to infer the presence, type, and intensity of divided attention during learning.

RELATED WORK

By definition, divided attention (DA) occurs when attention is divided among simultaneous stimuli [8]. DA is different from mind wandering (MW) in that DA is either caused by *intentional* multi-tasking (internal distractions) or *passive* external distractions [8], while MW is an *involuntary* shift in attention from task-related thoughts to task-unrelated thoughts [11] and is *stimulus independent* [10]. Existing DA research in learning technology is limited because MW can happen in any environment [10] while DA is more pervasive in informal learning.

Research on DA has been focusing on understanding human capabilities to perform multiple tasks simultaneously [8, 4, 7]. Kahneman [8] systematically reviewed experiments on the parallel processing of simultaneous inputs and found that although parallel processing was possible, its effectiveness was often impaired due to the interferences among multiple activities. Craik et al. [4] conducted four experiments to explore the effects of DA on encoding and retrieval processes. Experimental results showed that DA during the encoding process was associated with large reductions in memory performance. The divided attention made the selection of information imperfect, resulting in delayed or slowed processes [15].

Researchers have explored the use of various physiological signals, such as Electroencephalography (EEG) [13, 16], eye gaze [1], galvanic skin responses (GSR) [2], and heart rate [12] to infer learners' attentional states in educational settings. However, most existing work focused on detecting MW or zoning out rather than divided attention.

The work by Rodrigue et al. [13] is perhaps the most relevant research. In two three-participant experiments, Rodrigue and colleagues built user-dependent models (accuracies range from 79% to 99%, depending on the user) to detect DA from signals collected by a consumer-grade eye tracker and an EEG sensor. OneMind advances the state-of-the-art by eliminating the requirements of dedicated sensors, investigating learning activities in mobile MOOC settings, and quantifying the impact of DA and its intensity on learning outcomes *statistically*.

THE DESIGN OF ONEMIND

OneMind adopts the *tangible video control interface* of AttentiveLearner [12, 17]. The back camera lens of mobile phones is used as the "play" button for video/media control. To play an instructional video, a learner uses his/her finger to cover and hold the camera lens. Uncovering the lens pauses the video. Detection of the lens covering actions is based on the *Static LensGesture* detection algorithm in [18]. Such a tangible video control integrates the lens-covering requirement in commodity camera based PPG sensing [6] (i.e., covering the camera lens with a fingertip to measure heart rate) naturally into video watching, thus allowing *implicit* physiological signal monitoring on unmodified mobile phones. Please refer to [17] for the implementation and evaluation details of the tangible video control interface.

While a learner is watching the lecture video, OneMind monitors her heart rate *implicitly* by analyzing the fingertip transparency changes captured by the back camera (commodity camera based PPG sensing) [6]. LivePulse algorithm [6] is used to extract learners' PPG signals and measure their heart rate. This algorithm itself is accurate and robust against motions of the device [6]. We used the commodity camera based PPG sensing, instead of a dedicated heart rate monitor, to capture heart rate signals for two reasons: 1) it works directly on unmodified smartphones and requires no extra devices, and 2) covering the camera lens to watch the videos can make learners pay more attention to the lecture [17].

USER STUDY

We ran an 18-participant study to investigate the impact of divided attention on both learning outcomes and learners' PPG signals. We studied two types of distractions: 1) multitasking distractions (i.e. *internal divided attention*) where the subject's attention is divided between two stimuli; and 2) unpredictable and intrusive auditory distractions (i.e. *external divided attention*).

Task

Each participant watched four lecture videos (8 minutes each) with OneMind. We used two types of stimuli to create internal divided attention condition (e.g. multitasking) and external divided attention condition (e.g. distractive audio sound) while participants were watching the MOOC videos. In the control condition (full attention), participants can focus on the video without any internal or external stimuli.

We adopted the color counting task [13] to introduce internal divided attention. While a participant was watching the lecture video using OneMind, a computer placed on her side spoke the names of six different colors in a random order at the speed of one-second per word (high divided attention) or five-second per word (low divided attention). Participants were told to focus on the video but also count the number of times a target color (e.g., the color "red") was spoken during the video. The participants reported the counted number of the target color after each video, which was compared to the ground truth to ensure that they indeed divided their attention between the two tasks.

To simulate an environment where a learner is distracted by external stimuli (external divided attention), such as unexpected and intrusive auditory distractions, the computer placed on the side of the learner played loud and energetic music while the learner was watching the lecture video. We chose to use music as an external stimulus because it is a common environmental sound during informal learning.

Although stimulus of the internal divided attention condition (the "color counting" task) also comes from an external source, it aims to create an internal distracted state, i.e. multitasking, when the learner is mentally performing multiple tasks at the same time.

Participants and Apparatus

Eighteen subjects (6 females) participated in the study (Figure 2). The average age was 25.8 (σ = 2.73) ranging from 22 to 32. All participants were graduate students from a local university. None of them had prior knowledge of the learning materials used in the study. Four subjects had prior experience of using mobile apps for MOOC learning.

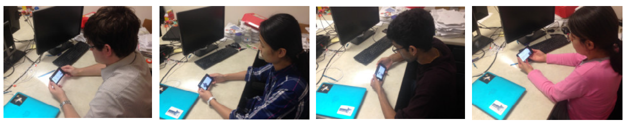

Figure 2. Some participants in the user study.

The experiment was completed on a Nexus 5 smartphone with a 4.95 inch, 1920 x 1080 pixel display.

Procedure

The study consisted of three parts:

Introduction: we ran a tutorial session and collected background information from participants.

MOOC Learning Session: Participants used OneMind to watch four video clips in four conditions (i.e. within-subjects design): control condition (i.e. Full Attention, **FA**), low internal divided attention (**LIDA**), high internal divided attention (**HIDA**), and external divided attention (**EDA**). The order of conditions assigned to each video was counterbalanced by a Latin Square pattern.

Two video clips were from the course "*Intro to Design of Everyday Things*", and the other two were from the course "*Intro to iOS App Development with Swift*". Both courses were taken from the MOOC learning platform, Udacity. All clips were edited to exact eight-minute long.

Post-video Quiz and Self-Rating: After finishing each video clip, participants reported the perceived distraction level for the current learning condition on a 7-point Likert scale. Participants also completed a five-question quiz (short answer questions) on the topics covered in the video.

RESULTS AND DISCUSSIONS

Perceived Distractions

The perceived distractions were 1.33 (σ = 0.47), 3.11 (σ = 1.59), 4.11 (σ = 1.24), and 4.97 (σ = 1.18) for the four conditions **FA**, **EDA**, **LIDA**, and **HIDA** respectively.

Repeated Measures of Analysis of Variance showed a significant main effect ($F(3, 15) = 9.28$, $p < 0.0001$) of the perceived distractions among the four conditions. Pairwise mean comparison (t-tests) with Bonferroni correction showed that the control condition (**FA**) was rated significantly less distractive than all other conditions ($t(17) = 4.74$, $p < 0.001$; $t(17) = 8.71$, $p < 0.001$; $t(17) = 12.27$, $p < 0.001$). *External Divided Attention (***EDA***)* was significantly less distractive than the *high internal divided attention (***HIDA***)* ($t(17) = 4.28$, $p < 0.001$). The difference between the two levels of internal divided attentions, i.e. **HIDA** and **LIDA,** was also significant ($t(17) = 3.38$, $p < 0.005$).

Effect of Divided Attention on Learning Performance

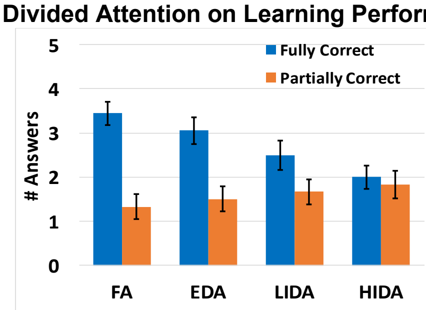

Figure 3. Participants' quiz performance by conditions (Error bars represent standard error of the mean).

Questions in the post-video quiz were graded by the following rubrics: a participant received 1 point for a complete and accurate answer; the participant received 0.5 points for an answer with a correct general idea but missing critical details; the participant didn't receive any point if the answer was incorrect or missing. The grader was blind to the conditions when grading. Figure 3 shows the distributions of correct and partially correct answers by conditions.

Participants' average scores for the four conditions were 4.11 (σ = 0.66), 3.83 (σ = 0.90), 3.33 (σ = 0.92), and 2.92 (σ = 0.98) respectively. Repeated measures of ANOVA showed a significant main effect ($F(3, 15) = 1.12$, $p < 0.01$) of the learning outcomes. Pair-wise mean comparison (t-tests) with Bonferroni correction showed that participants performed significantly better in *control condition* (**FA**) than both **LIDA** ($t(17) = -3.0$, $p < 0.01$) and **HIDA** ($t(17) = -4.13$, $p < 0.001$). Although the learners' average performance in **FA** was better than that in **EDA**, the difference was not significant ($t(17) = -1.17$, $p = 0.256$). Participants also performed significantly better in **EDA** than in **HIDA** ($t(17) = -3.41$, $p < 0.005$). These results suggest that **IDA** is more detrimental to learning than **EDA**.

Compared to the full attention condition, participants had less entirely correct answers, but more partially correct answers in the divided attention conditions. This is especially true for the **HIDA** condition, where the number of entirely correct answers were only 58.1% of that in the **FA** condition (2 vs. 3.44), while the number of partially correct answers increased by 37.5% (1.83 vs. 1.33). We found that learners in the divided attention condition were more likely to miss important details. Their answers also showed partial and shallower understanding of the learning materials.

Detecting Divided Attention

Considering the negative impact of divided attention on learning in MOOCs, we explored the use of PPG signals *implicitly* captured by OneMind during the study to predict whether a learner has divided attention, as well as its type and intensity while they watched each video.

OneMind collected four PPG signal sequences from each participant (72 PPG signal sequences in total), one for each condition (**FA**, **EDA**, **LIDA**, **HIDA**). The PPG signal sequences were first resampled at 20Hz before going through a second-order Butterworth filter with a cutoff frequency of [0.75, 3.3] Hz to reduce noises.

Classification Task	Accuracy	Precision	Recall	Kappa
FA vs. EDA	72.2%	0.75	0.67	0.44
FA vs. LIDA	75.0%	0.71	0.83	0.50
FA vs. HIDA	83.3%	0.80	0.89	0.67
FA + EDA vs. LIDA + HIDA	83.3%	0.90	0.75	0.67
FA vs. LIDA vs. HIDA	63.0%	0.67	0.63	0.44
FA vs. EDA vs. LIDA vs. HIDA	50.0%	0.52	0.50	0.33

Table 1. Performance of the RBF-SVM classifiers.

We used the LivePulse algorithm [6] to extract RR-intervals (i.e., the cardiac interval between two adjacent heart beats) from each PPG signal sequence. Outliers of the RR-intervals are removed using the same heuristics in [17]. Eleven dimensions of heart rate variability (HRV, variation in the time interval between heart beat) features were extracted from each signal sequence: 1) AVNN (average RR-intervals); 2) SDNN (standard deviation of the RR-intervals); 3) rMSSD (square root of the mean of the squares of difference between adjacent RR-intervals); 4-7) pNN5, pNN10, pNN20, pNN50 (percentage of adjacent RR-intervals with a difference longer than 5ms, 10ms, 20ms, and 50ms); 8) MAD (median absolute deviation of all RR-intervals); 9) SDANN (standard deviation of AVNN in all k segments of a sequence), 10) SDNNIDX (mean of SDNN in all k segments of a sequence), and 11) rMSSD/SDNNIDX. These features were common short-term time domain HRV features [12, 14].

We used Weka to build classifiers which used the extracted HRV features to predict a learner's attentional states during the time of a given PPG signal sequence. We used leave-one-subject-out evaluation to train different classifiers. Therefore, our classifiers were user-independent. SVM with a RBF kernel yielded the best overall performance. Table 1 lists the classification performance achieved by the RBF-SVM classifier for various classification tasks (For example, **FA** vs. **EDA** means distinguishing between the **FA** samples and the **EDA** samples). The chance level accuracy for classification tasks in Table 1 were 50.0%, 50.0%, 50.0%, 50.0%, 33.3% and 25% respectively (kappa=0). It is expected that the detection accuracy of **HIDA** is higher than **LIDA** and **EDA**, considering that **HIDA** may cause stronger changes in physiological arousal due to the higher level of divided attention.

Both the accuracy and Kappa score of our classifiers were far-above chance, indicating the feasibility of using OneMind to infer the presence, type and intensity of divided attention in a user-independent fashion. We believe that higher accuracies could be achieved by training personalized, user-dependent classifiers. Moreover, given a moderate degree of recognition accuracy, we can focus on developing less intrusive, fail-soft attention-aware interventions (e.g. polite reminders or content review requests) that do no harm if delivered incorrectly.

We also ran a linear regression analysis to gain further insights on the relationship between HRV features and different attentional states. We found that for **FA** vs. **EDA** classification, the important features were MAD ($p = 0.0167$) and pNN5 ($p = 0.0258$); while important features for **FA** vs. **IDA** classification were MAD ($p = 0.0054$), SDNN ($p = 0.0158$), and SDANN ($p = 0.0326$). For **FA** vs. **EDA** vs. **LIDA** vs. **HIDA** classification, the only important feature was MAD ($p = 0.0325$). These results suggested that MAD was the most important features for predicting divided attention in MOOCs.

DISCUSSIONS

The automatically detected information on DA can be leveraged to improve mobile MOOC learning in at least two ways. First, since different types of learning activities have different demands for attention, the system could switch to less attention-demanding learning activities, such as discussion forums, when consistent DA is detected; Second, proper intervention technologies could be introduced to directly address DA. For example, a system can use visual and tactile feedback to remind learners to focus on the video when DA is detected.

Since OneMind does not require dedicated physiological sensing devices, it has the potential to reach millions of learners on today's MOOC platforms. However, it is worth noticing that the OneMind algorithm can also work on signals collected from dedicated physiological sensors or wearable devices. Moreover, other than its application in education, the idea of implicit PPG signal sensing and divided attention detection could be adopted in areas such as usability evaluation, adjusting the difficulty of games, and evaluating the quality of advertising.

CONCLUSIONS AND FUTURE WORK

We proposed OneMind, a novel system for detecting the presence of divided attention during mobile MOOC learning via implicit heart rate monitoring on today's smartphones without any hardware modification. In an 18-participant study, we observed negative impacts of divided attention on students' learning outcomes. Additionally, we systemically studied the use of PPG signals implicitly collected by OneMind during the learning process to predict the existence, type, and intensity of divided attention.

REFERENCES

1. Robert Bixler and Sidney D'Mello. 2014. Toward Fully Automated Person-Independent Detection of Mind Wandering. In *International Conference on User Modeling, Adaptation, and Personalization* (UMAP 2014), 37-48.
2. Nathaniel Blanchard, Robert Bixler, Tera Joyce, and Sidney D'Mello. 2014. Automated Physiological-Based Detection of Mind Wandering during Learning. In *International Conference on Intelligent Tutoring Systems* (ITS 2014), 55-60.
3. Phillip Compeau and Pavel A. Pevzner. 2015. Life After MOOCs. *Communications of the ACM* 58.10 (2015): 41-44.
4. Fergus IM Craik, Richard Govoni, Moshe Naveh-Benjamin, and Nicole D. Anderson. 1996. The Effects of Divided Attention on Encoding and Retrieval Processes in Human Memory. *Journal of Experimental Psychology: General* 125.2 (1996): 159.
5. Phillip J. Guo, Juho Kim, and Rob Rubin. 2014. How Video Production Affects Student Engagement: An Empirical Study of MOOC Videos. In *Proceedings of the first ACM conference on Learning@ scale conference*, 41-50
6. Teng Han, Xiang Xiao, Lanfei Shi, John Canny, and Jingtao Wang. 2015. Balancing Accuracy and Fun: Designing Engaging Camera Based Mobile Games for Implicit Heart Rate Monitoring, In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI 2015), 847-856.
7. Larry L. Jacoby. 1991. A Process Dissociation Framework: Separating Automatic from Intentional Uses of Memory. *Journal of memory and language* 30.5 (1991): 513-541.
8. Daniel Kahneman. 1973. *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
9. Thierry Karsenti. 2013. What the Research Says. *International Journal of Technologies in Higher Education* 10.2 (2013): 23-37.
10. Matthew A. Killingsworth and Daniel T. Gilbert. 2010. A Wandering Mind is an Unhappy Mind. *Science 330.6006* (2010): 932-932.
11. Malia F. Mason, Michael I. Norton, John D. Van Horn, Daniel M. Wegner, Scott T. Grafton, and C. Neil Macrae. 2007. Wandering Minds: the Default Network and Stimulus-Independent Thought. *Science 315.5810* (2007): 393-395.
12. Phuong Pham and Jingtao Wang. 2015. AttentiveLearner: Improving Mobile MOOC Learning via Implicit Heart Rate Tracking. In *International Conference on Artificial Intelligence in Education* (AIED 2015), 367-376.
13. Mathieu Rodrigue, Jungah Son, Barry Giesbrecht, Matthew Turk, and Tobias Höllerer. 2015. Spatio-Temporal Detection of Divided Attention in Reading Applications Using EEG and Eye Tracking. In *Proceedings of the International Conference on Intelligent User Interfaces* (IUI 2015), 121-125.
14. Kristina Schaaff and Marc TP Adam. 2013. Measuring Emotional Arousal for Online Applications: Evaluation of Ultra-Short Term Heart Rate Variability Measures. In *Affective Computing and Intelligent Interaction* (ACII 2013), 362-368.
15. Edward E. Smith and Stephen M. Kosslyn. 2013. *Cognitive Psychology: Pearson New International Edition: Mind and Brain.* Pearson Higher Ed.
16. Daniel Szafir and Bilge Mutlu. 2013. ARTFul: Adaptive Review Technology for Flipped Learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI 2013), 1001-1010.
17. Xiang Xiao and Jingtao Wang. 2015. Towards Attentive, Bi-directional MOOC Learning on Mobile Devices. In *Proceedings of the International Conference on Multimodal Interaction* (ICMI 2015), 163-170.
18. Xiang Xiao and Jingtao Wang. 2013. LensGesture: Augmenting Mobile Interactions with Back-of-Device Finger Gestures. In *Proceedings of the International Conference on Multimodal Interaction* (ICMI 2013), 287-294.