

Old Dominion University

ODU Digital Commons

Communication Disorders & Special Education
Faculty Publications

Communication Disorders & Special Education

2019

Predicting ADHD Using Eye Gaze Metrics Indexing Working Memory Capacity

Anne M.P. Michalek

Old Dominion University, aperrott@odu.edu

Gavindya Jayawardena

Old Dominion University, hjayawar@odu.edu

Sampath Jayarathna

Old Dominion University, ujayarat@odu.edu

Follow this and additional works at: https://digitalcommons.odu.edu/cdse_pubs



Part of the [Computer Sciences Commons](#), and the [Educational Psychology Commons](#)

Original Publication Citation


Michalek, A. M. P., Jayawardena, G., & Jayarathna, S. (2019). Predicting ADHD using eye gaze metrics indexing working memory capacity. In *Computational Models for Biomedical Reasoning and Problem Solving* (pp. 66-88). Hershey, PA: IGI Global.

This Book Chapter is brought to you for free and open access by the Communication Disorders & Special Education at ODU Digital Commons. It has been accepted for inclusion in Communication Disorders & Special Education Faculty Publications by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.

Chapter 3

Predicting ADHD Using Eye Gaze Metrics Indexing Working Memory Capacity

Anne M. P. Michalek

 <https://orcid.org/0000-0001-6850-3948>
Old Dominion University, USA

Gavindya Jayawardena

Old Dominion University, USA

Sampath Jayarathna

Old Dominion University, USA

ABSTRACT

ADHD is being recognized as a diagnosis that persists into adulthood impacting educational and economic outcomes. There is an increased need to accurately diagnose this population through the development of reliable and valid outcome measures reflecting core diagnostic criteria. For example, adults with ADHD have reduced working memory capacity (WMC) when compared to their peers. A reduction in WMC indicates attention control deficits which align with many symptoms outlined on behavioral checklists used to diagnose ADHD. Using computational methods, such as machine learning, to generate a relationship between ADHD and measures of WMC would be useful to advancing our understanding and treatment of ADHD in adults. This chapter will outline a feasibility study in which eye tracking was used to measure eye gaze metrics during a WMC task for adults with and without ADHD and machine learning algorithms were applied to generate a feature set unique to the ADHD diagnosis. The chapter will summarize the purpose, methods, results, and impact of this study.

DOI: 10.4018/978-1-5225-7467-5.ch003

INTRODUCTION

Attention-Deficit/Hyperactivity Disorder is being recognized as a diagnosis which persists into adulthood impacting economic, occupational, and educational outcomes. Estimates indicate that 3-5% of adults have a diagnosis of ADHD (Willcutt 2012) with prevalence estimated to have increased from 6.1% of the United States population in 1997 to 10.2% of the population in 2016 (Xu et al. 2018). The disorder is behaviorally marked by difficulty with attention to important details, difficulty initiating and completing tasks, and difficulty modulating behaviors appropriately in relation to the situation (Fields et al. 2017; Fostick 2017). According to Barkley (1997), adult ADHD symptoms result from impairments of inhibition or the inability to regulate and modulate prepotent responses. While a diagnosis of adult ADHD presumes disinhibition, little is known about the physiological underpinnings of that cognitive skill in relation to an adult ADHD diagnosis. There is an increased need to accurately diagnose ADHD through the development and implementation of objective and reliable outcome measures which reflect core diagnostic criteria, like inhibition.

Researchers in cognitive psychology evidence attention control as the measurable psychological construct which facilitates inhibitory responses by allocating attention according to task demands, especially in the presence of distracting stimuli (Conway et al. 2005; Engle 2002; Kane et al. 2001). Attention control differentiates success during tasks requiring intentional and sustained constraints for effective inhibition, like dichotic listening (Colflesh and Conway 2007) or processing speech in noise (Rönnerberg et al. 2013). Measurements of attention control are demonstrated through differences in working memory capacity (WMC) accounting for approximately 60% of the variance seen across people on measures of WMC, like complex span tasks (Engle et al. 1999). Adults with ADHD have reduced WMC when compared to their peers (Michalek et al. 2014) and, despite the understanding that disinhibition is central to an ADHD diagnosis and differences in WMC mathematically represent the resource which makes inhibition possible, there is a paucity of research investigating physiological responses during measures of WMC which could differentiate adults with and without ADHD.

The primary goal of this work is to determine the feasibility of identifying and integrating eye gaze metrics from a WMC task using machine learning to generate a valid and reliable feature set which indexes and predicts an ADHD

diagnosis. This chapter investigates gaze measures that map onto these valid neurocognitive deficits that are central to ADHD within the context of a WMC task. The development of these objective measures of ADHD will facilitate its diagnosis and reveal strategies that can enhance the future design effective intervention strategies and accessible classroom environments.

BACKGROUND

Working Memory Capacity (WMC) Tasks

Working memory is the cognitive system which makes it possible to mentally hold and manipulate information simultaneously. Over a decade of work by Engle and colleagues Conway et al. (2008) supports the use of complex span tasks as not only measures of working memory but as a reflection of individual differences in WMC. Performance on complex span tasks generates a composite working memory score numerically indicating how well someone can manipulate and hold information. However, differences in WMC or that composite score represent a person's ability to moderate and control attention. Adults with ADHD have reduced working memory when compared to their peers demonstrating significant differences in WMC (Alderson et al. 2013). This finding suggests that adults with ADHD have a reduced ability to monitor and control attention, especially during situations with competing stimuli (Michalek et al. 2014) or that require response inhibition (Lee et al. 2015; Roberts et al. 2011). Little is known about the underlying covert processes engaged during inhibitory tasks which rely on attention allocation. Using physiological measures during a task which validly reflects attention control, like a complex span task, provides objective diagnostic information for adults with ADHD.

The reading span task (R-Span) is one complex span task widely used as a valid measure of working memory yielding a WMC score (Conway et al. 2005). The R-span was originally developed by Daneman and Carpenter (1983) as a predictor of reading comprehension and was subsequently modified by Engle and colleagues Engle et al. (1999). During the R-Span people see one sentence on a computer screen, read the sentence out loud, determine if the sentence is meaningful with a yes or no response, and then verbally identify a letter typed at the end of each sentence. After a set of sentences, the person is asked to verbally recall as many letters as possible in order of presentation.

This task represents the person's ability to hold and manipulate information simultaneously. To date, there have been no empirical studies investigating eye gaze metrics collected during this task which might differentiate performance and further explain diagnostic differences for adults with and without ADHD.

Machine Learning and ADHD

Typically, experts diagnose ADHD using subjective checklists and related academic and cognitive performance measures (Greenhill 1998). These comprehensive assessments can be time consuming, inconsistent, and can inaccurately represent deficits making differential diagnosis challenging. Ideally, it would be more efficient and reliable to develop predictive algorithms based on physiological metrics reflecting core diagnostic criteria. However, designing this type of computer program can be difficult because there is no existing set of confirmed mathematical features which accurately differentiate between adults with and without ADHD. Machine learning principles offer a solution to this barrier. While it is not practical to develop algorithms by providing a specific set of instructions, machine learning uses numeric features representing a cognitive skill to teach the computer data patterns and inferences which can be applied to groups of data for predicting accurate classifications.

In machine learning there are several varieties of sub categorical learning algorithms. Supervised learning is an example of such a subcategory. The core objective of supervised learning is to build a mathematical model which can be used to predict the outputs of new samples using the training data. Usually, the training data set is stored in a matrix. Each row of the training data matrix corresponds to one training instance which also contains the desired output. For example, in the ADHD/Non-ADHD literature, supervised learning algorithms are iterated through the training dataset to learn a mathematical model to predict or classify the output associated with unseen inputs. When determining whether a person does or does not have ADHD using eye gaze as the outcome metric, the training data would consist of eye tracking data of each person and each person would have a class specifying whether that person is identified as having ADHD or not having ADHD. Supervised learning algorithm will build a general mathematical model which covers the training data space. When a previously unseen person's eye gaze data is entered, the model will use its past experience to accurately predict whether or not that person has ADHD.

Some experts may question the reliability and validity of using machine learning for predicted outputs because those outputs do not consider experts knowledge from psychology or medicine in the decision process. Even with the involvement of expert physicians, it is reported that “diagnostic errors contribute to approximately 10% of patient deaths”, by Institute of Medicine at the National Academies of Science, Engineering and Medicine (National Academies of Sciences et al. 2016). Causes for such diagnostic errors could be communication errors between patients and physicians and other failures of the healthcare system. These challenges could be addressed by identifying patterns of the symptoms patients confirm and use them to predict potential diagnostic codes. Even if there is a lack of communication between the patient and the physician or there is a failure in a healthcare system, the symptom patterns of the patient would be highlighted so that accurate diagnosis prediction could be facilitated.

Currently, machine learning is being used for diagnostic prediction not only based on reported symptoms, but also based on patient history and data extracted from wearable devices. Classification algorithms compare the symptoms pattern of the patient and other related data with the other patients in the training dataset in order to predict an accurate diagnosis. In literature using machine learning for ADHD diagnosis and classification, Mueller et al. (2010) attempt to classify adult ADHD patients and healthy controls using a machine learning algorithms. They conducted research to classify ADHD patients and healthy controls using support vector machine (SVM) learning based on event related potential (ERP) components. They examined data from 148 adult participants. Among them, 50% were diagnosed as ADHD while the rest did not have a diagnosis of ADHD. Both groups of adults were selected in a manner that age and gender did not vary between the two group. Each participant performed a visual two stimulus GO/NOGO task and ERP responses of participants were decomposed into independent components and created the feature set. Classification accuracy of assigning ADHD participants and healthy controls to the corresponding groups using a non-linear SVM with 10-fold cross-validation was 92%, whereas it was 90% for linear SVM. This research suggests that classification by means of non-linear methods is more accurate for experiments conducted in a clinical context. Even-though this study uses machine learning approaches to predict ADHD, it does not use eye gaze metrics nor working memory capacity.

Peng et al. (2013) shows that extreme learning machine (ELM), a machine learning algorithm, achieves 90.18% accuracy when predicting ADHD using structural MRI data. This study confirmed that linear support vector

machine and support vector machine-RBF achieves an accuracy of 84.73% and 86.55% respectively when the same structural MRI dataset is used. Both extreme learning machine and support vector machine have been evaluated to find the classification accuracy using cross-validation. The goal of their study was proposing an ADHD classification model using the extreme learning machine (ELM) algorithm for ADHD diagnosis. They assessed the computational efficiency and the effect of sample size on both extreme learning machine and support vector machine. They acquired MRI images from 110 participants with 50% of them having a diagnosis of ADHD. This study gives us insight about how applying a machine learning model can accurately predict ADHD.

Marcano et al. (2016) aimed to classify people with ADHD and without ADHD using autoregressive models. They used EEG data collected using 26 electrodes from a group of children between the ages of 6 and 8 to discriminate between ADHD and Non-ADHD. Children participated in multiple experimental conditions, such as eyes open, eyes closed, and quiet video baseline tasks while collecting EEG data. This study verified that KNN classifier is able to provide high classification accuracy when classifying children as either ADHD or typically developing ADHD. The accuracy achieved in this study was high and varying between 85% and 95%. Abibullaev and An (2012) used EEG data with semi-supervised learning in order to predict a ADHD and Non-ADHD diagnosis. They had 10 children participants with 7 of them having a diagnosis of ADHD and 3 were typically. They trained and tested support vector machine with EEG data of each participant achieving an accuracy of 97% for ADHD prediction using support vector machine learning.

Taken together, the literature affirms the successful use of machine learning to accurately predicate a diagnosis of ADHD. However, an empirical gap exists with regard to the training dataset used to train the machine learning model. The majority of studies conducted have primarily used MRI, fMRI, or EEG data to train the machine learning model used to discriminate between ADHD and Non-ADHD. Our goal of this work is to predict diagnosis of ADHD using eye gaze metrics and measures of working memory as the training data for the machine learning algorithms. In this work we employ machine learning algorithms for a diagnosis of ADHD using eye gaze metrics collected during a WMC task. The output of our task is based on binary classes, ADHD and Non-ADHD. We employed supervised learning classification algorithms and evaluated the predicted output in terms of how close it is to the actual output. Our evaluation matrices include standard information retrieval domain evaluation measurements such as accuracy, precision, recall, and f1.

Eye Movements and ADHD

Eye movement behavior is a result of complex cognitive processes; therefore, eye gaze metrics can reveal objective and quantifiable information about the quality, predictability, and consistency of these covert processes (Van der Stigchel et al. 2007). Eye gaze measurement includes a number of metrics relevant to oculomotor control (Komogortsev et al. 2013) including saccadic trajectories, fixations, and other relevant measures - such as velocity, duration, amplitude, pupil dilation (Krejtz et al. 2018). A saccade (rapid eye movement from one fixation point to another) itself may not be an informative indicator of cognition since visual perception is suppressed during a saccade. However, fixations require preceding saccades to help place the gaze on target stimuli to gather salient and relevant information. We believe that analysis of these eye movements can provide important cumulative clues about the underlying physiological functions of attention control during a WMC task which can differentiate a diagnosis of ADHD for adults.

There is substantial overlap in brain systems that are involved in oculomotor control and cognitive dysfunction in ADHD. The precise measurements of eye movements during cognitively demanding tasks provide a window into underlying brain systems affected by ADHD. The neural substrates of oculomotor control are well established (Leigh and Kennard 2004) and show proximity to and overlap with the cortical and subcortical structures involved in cognitive dysfunction in ADHD. For example, the cortical structures that mediate saccadic programming as well as a number of saccadic behaviors include frontal-parietal areas such the frontal eye field (FEF), supplementary eye field (SEF), parietal eye field (PEF), and DLPC. These areas are also affected during cognitive control and WM in ADHD (Rubia 2018). With respect to subcortical structures, the accuracy of saccades is maintained via cerebellum. For example, saccadic hypometria is an undershooting of a saccade to a target that is typically seen in normal subjects, whereas saccadic hypermetria, overshooting the target, is a hallmark feature of cerebellar dysfunction (Leigh and Zee 2015). A study of saccades during visuo-spatial WM has reported significant diagnostic group differences in under- versus over-shooting to the target between boys with ADHD and non-affected controls, such that the ADHD group tended to overshoot the target and the control group tended undershoot the target (Rommelse et al. 2008). Studies have shown that individuals with ADHD also have deficits in the suppression of saccades relative to controls (Mostofsky et al. 2001; Munoz et al. 2003;

Rommelse et al. 2008). Similarly, people with ADHD demonstrate difficulties with intentionally inhibiting ocular responses when compared to their peers during tasks which require purposeful anti-saccade behaviors (Lee et al. 2015; Roberts et al. 2011). Eye gaze metrics, especially saccade features, reliably reveal important differences between adults with and without ADHD.

Based on the diagnostic utility of eye movements, Blazey et al. (2003) invented a method which determines whether an individual has ADHD by sampling the eye movements of participants when they are in an inactive state. This procedure includes a sampling device which has infrared radiation for brightening the eye of a participant and detecting reflections from the eye. The eye movement data collected using their device determines the value of a pre-selected parameter which has a threshold value indicating whether the participant has ADHD or not. According to their study, the most significant feature of the eye movement data is the angular acceleration of the eyeball (Blazey et al. 2003). They have measured ocular angular acceleration for the participants by asking them to stare at a blank screen (Blazey et al. 2003). The measurement data of the angular acceleration of the eye below the threshold value indicates diagnosis of ADHD and the data above the threshold value constitutes a classification of healthy/normal.

Eye Gaze and Machine Learning and ADHD

There is a paucity of empirical studies which implement machine learning to predict ADHD classification using a measures of attention control or WMC. However, there are a few investigations which use machine learning in combination with measures of inhibition. Hart et al. (2014) measured activation patterns using functional magnetic resonance imaging while adolescents with ADHD performed a Stop Task. During this task, participants had to suppress or inhibit the motoric response of pushing a button. The researchers used Gaussian process classifiers and whole activation pattern analysis and were able to predict the ADHD diagnosis with 77% accuracy. Likewise, in a study with adults with and without a diagnosis of ADHD, machine learning predicted the diagnosis with a specificity of .91 and sensitivity of .76 based on EEG metrics during a NoGo task measuring inhibition (Biederman et al. 2017). These results support collecting physiological metrics during tasks required attention control to generate pattern recognition analysis for the accurate classification of ADHD.

Similarly (Tseng et al. 2013) used eye movements in conjunction with machine learning to predict ADHD. This study included participants diagnosed with ADHD, fetal alcohol spectrum disorder (FASD), and Parkinson's disease (PD). Researchers presented short video clips to each participant and analyzed the resulting data sets for three specific types of eye movement features: 1) oculomotor-based features such as fixation durations and distributions of saccade amplitudes; 2) saliency-based features; and 3) group-based features. Results confirmed that saliency based features best differentiated children with ADHD and FASD from typically developing children. Machine learning algorithms predicted ADHD in the sample of children with 77.3% accuracy. Taken together, these empirical findings suggest that diagnostic biomarkers of ADHD could be generated from eye gaze metrics during a WMC task using machine learning.

As such, in this feasibility study, we examined patterns of saccades and stability of fixations generated when completing a measure of WMC to create a feature set which could be used to differentiate a diagnosis of ADHD for adults. Based on the evidence that WMC is reduced in adults with ADHD (Michalek et al. 2014), measurement of eye movements during a measure of WMC will address the following research question: 1) do eye gaze feature values indexing a WMC task predict the classification of ADHD in adults?

METHODOLOGY

Participants

A total of 14 adult participants without ($n = 7$) and with a diagnosis ($n = 7$) between the ages of 18-35 were recruited for this study from Old Dominion University. The seven adult participants were (6 F, 1 M, $M_{age}=22.85$, $SD_{age}=3.01$) diagnosed with ADHD by medical practitioners and that diagnosis was confirmed through formal and verified documentation. Each ADHD participant also completed an informational interview verbally confirming their diagnosis. Moreover, adults with ADHD remained medication free for the 12 hours prior to study participation. Prior to beginning study tasks, all adults were informed of their risks regarding remaining medication free and participating in the study. Participants provided their consent by signing forms outlining costs and benefits of participation approved the University's Institutional Review Board (IRB) in accordance with the Helsinki Declaration.

Participants who completed the protocol were given a ten-dollar Amazon or Chick-Fil-A gift card.

Working Memory Capacity Task

WMC is reflected through complex span tasks, including the Reading Span (R-Span). The R-Span is a validated task designed to reflect the cognitive system's ability to maintain activated representations (Engle et al. 1999; Engle 2002). In the R-Span task, participants are asked to read a sentence and letter they see on a computer screen. Sentences are presented in varying sets of 2-5 sentences. Participants are asked to judge sentence coherency by saying 'yes' or 'no' at the end of each sentence. Then, participants are asked to remember the letter printed at the end of the sentence. After a 2-5 sentence set, participants are asked to recall all the letters they can remember from that set. WMC scores are generated based on the number of letters accurately recalled divided by the total number of possible letters recalled in order. However, this work focused on measures of visual attention which could differentiate adults with and without ADHD.

Apparatus

Eye gaze metrics were recorded and analyzed using the Tobii Pro X2-60 computer screen-based eye tracker with Tobii Studio analysis software. The Tobii Pro X2-60 records eye movements using infrared corneal reflective technology at a sampling rate of 60 Hz (i.e. approximately once every 16.23 milliseconds). Gaze data accuracy was within 0.4 degrees of visual angle and precision was within 0.34 degrees of visual angle. Tobii's eye tracking technology is effective for generating reliable and valid brain/behavior outcomes for children and adolescents (Richmond and Nelson 2009).

All of the participants fulfilled the following inclusion criteria:

1. Between 18 and 65 years of age.
2. Spoke English as their first language.
3. Self-reported normal vision with or without corrective lenses.
4. No history of psychotic symptoms.
5. No comorbid cognitive impairments (e.g. documented learning disabilities, reading disabilities).

Eye Movement Features

The human oculomotor plant (OP) (Komogortsev et al. 2010) consists of the eye globe and six extraocular muscles and its surrounding tissues, ligaments each containing thick and thin filaments, tendon-like components and liquids. In general, there are six major eye movement types: fixations, saccades, smooth pursuits, optokinetic reflex, vestibule-ocular reflex and vergence (Leigh and Zee 2015). An eye-tracker provides eye gaze position information as well as other gaze related parameters (pupil dilation etc.) so that algorithmic derivation in terms of two primary eye movements, fixations (relative gaze position at one point on the screen) and saccades (rapid eye movements of gaze from one fixation point to another) can be analyzed to derive the users attention patterns.

We are interested in investigating number of eye fixation based features in the current framework. We developed a detailed saccade and fixation feature set using the following qualifiers: gender, number of fixations, fixation duration measured in milliseconds, average fixation duration in milliseconds, fixation standard deviation in milliseconds, pupil diameter left, pupil diameter right, and diagnosis label or class. Due to the sampling rate of the tracking system, we were not able to calculate microsaccades and overshoot/undershoot saccades as components of the feature set.

Measuring Attention During WMC

The data for this study was collected during a larger project involving adults with and without ADHD and an audiovisual listening in noise task where WMC scores were measured and used as a cognitive covariate without eye tracking metrics. The entire testing session for the study took approximately 45 minutes. The session began with the participant interview, explanation of the purpose of the study, and review of the consent form. During the interview, the participants provided demographic information and were screened to confirm that all inclusion criteria were satisfied. Once participants indicated they understood their rights and gave consent, they entered the testing area to begin the study. Participants sat at a desk in front of a Dell Computer with a 21-inch monitor. The distance and position of each participant was modified in order to maintain a 45 degree viewing angel of the monitor. For each participant, the experimental tasks began with eye gaze calibration. Once calibration was confirmed, participants viewed a welcome screen followed

by the random presentation of several experimental tasks, including the R-SPAN task. The location of the R-SPAN in the order of experimental tasks was randomized and counterbalanced across participants in order to maintain validity. Participants were randomly assigned to a group determining the order of experimental task presentation prior to beginning the study. For all of the experimental tasks, participants were given practice trials.

MACHINE LEARNING ON DATA

We chose precision, recall, f-measure, and accuracy as the evaluation measures for our work. Prior studies (Manevitz and Yousef 2007) have already proven that these measures are independent of category distributions provided that precision and recall are measured at the same time. Intuitively, precision measures exactness of the system (i.e., out of all predicted data instances for a specific category label how many are predicted correctly) while recall indicates the completeness of the system (i.e., out of all labeled data for a specific a category label how many are predicted correctly). F value measures the balance between precision and recall in a single value. In our tables with results assessing classifiers, precision, and recall refers to their weighted average values. Accuracy specifies the fraction of the predictions that the classifier predicated correctly. We employed a grid search mechanism to identify the best parameter combination for optimal result. The optimal parameters are selected based on performance for each classifier after a 10-fold cross validation.

RESULTS

Table 1 shows the RSPAN score for the participants in the current study. An independent t-test statistical analysis ($p=0.07$) confirms that for this feasibility study there are no significant group differences on WMC scores and are predicted to be a result of the small sample size. The RSPAN is an individual differences measure and significance in variance is detected with large sample sizes. Additionally, WMC scores are typically generated through a composite score of two or more span tasks (Conway et al. 2005), for example, a previous investigation by one of the authors confirmed group differences in WMC using the RSPAN and operation span (OSPAN) to generate a WMC composite score for adults with ADHD (Michalek et al. 2014).

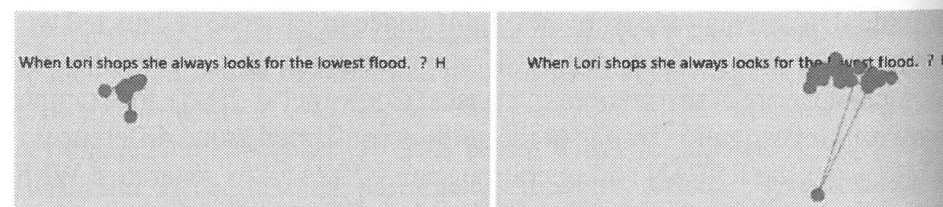
Table 1. RSPAN Score of the ADHD Vs Non-ADHD.

Participant	Age	Gender	RSPAN	Classification
3	18	Female	0.86	Non-ADHD
7	35	Male	0.88	Non-ADHD
9	19	Female	0.60	Non-ADHD
17	23	Male	0.55	Non-ADHD
20	21	Female	0.57	Non-ADHD
25	32	Male	0.88	Non-ADHD
26	20	Female	0.74	Non-ADHD
30	21	Female	0.51	ADHD
34	19	Male	0.67	ADHD
35	26	Female	0.76	ADHD
36	29	Female	0.71	ADHD
37	21	Female	0.60	ADHD
38	21	Female	0.40	ADHD
47	23	Female	0.62	ADHD

Visual Analysis

Figure 1 presents images of eye gaze patterns from two adult participants, one with and one without ADHD. Informal visual analysis indicates that the adult with ADHD is fixating primarily below the stimulus items with little direct fixation to sentence components including: the words, the decision point, or the item to be remembered. Unlike the adults with ADHD, the adult without ADHD has a majority of fixations which are in-line with all sentence components. Although this is a conclusion generated from informal visual inspection, it reveals that adults with ADHD are not visually scanning stimulus

Figure 1. Comparison of Eye Fixations for ADHD (Left) and Non-ADHD (Right) participant during WMC Task.



items in a path similar to adults without ADHD. The fixation cluster pattern is just below the stimulus sentence components. This is consistent with the findings of Krejtz et al. (2015) who suggest that while adults with ADHD had similar fixations to salient visual cues when compared to adults without ADHD, they demonstrated less structured and more chaotic scan patterns.

Machine Learning for Classification Prediction

We generated three feature sets categorized according to metric type:

1. Fixation feature set.
2. Saccade feature set.
3. Saccade and Fixation combination feature set.

Each of the three feature sets were individually entered into 43 different classifiers yielding precision rates, recall rates, F1 scores, and percent accuracy. We identified six of the top performing classifiers for each of the three feature sets: J48, LMT, RandomForest, REPTree, K Star, and Bagging. Results for each feature set are discussed individually.

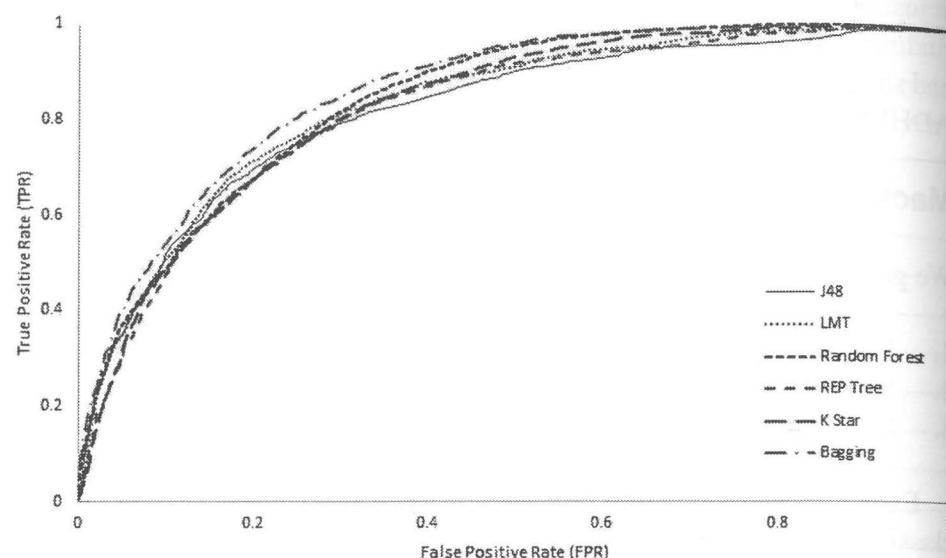
Six of the top performing classifiers for the fixation feature set are listed in the Table 2. The Bagging classifier (ensemble meta-estimator) yielded the highest percent accuracy with 78.48% indicating that a fixation feature set collected during a RSPAN task classifies a diagnosis of ADHD with greater than 70% accuracy. The REPTree classifier yielded the lowest percent accuracy at 76.77%.

To further investigate the performance metrics for the 6 most effective classifiers for the fixation feature set we generated a Receiver Operating Characteristics (ROC) graph (see Figure 2). The ROC graph displays the

Table 2. Classification of Eye Fixation Features during WMC.

Classifier	Precision	Recall	F1	Accuracy
J48	0.77	0.76	0.77	77.79
LMT	0.77	0.77	0.77	77.92
RandomForest	0.76	0.76	0.76	76.79
REPTree	0.75	0.76	0.75	76.77
K*	0.76	0.76	0.76	76.92
Bagging	0.77	0.78	0.77	78.48

Figure 2. ROC Graph of the Top Performing Classifiers for Fixation Feature Set



*For a more accurate representation see the electronic version.

relative trade-off between benefits (true positive) rates on the Y axis and the costs (false positive) rate on the X axis. The graph shows the Bagging as our top performing classifier offering the best trade-off in terms of the cost and the benefits.

Table 3 outlines six of the top performing classifiers for the saccade feature set. The Random Forest classifier yielded the highest percent accuracy at 91.14% indicating that the saccade feature set collected during a RSPAN task classifies a diagnosis of ADHD with greater than 90% accuracy. The J48 classifier yielded the lowest percent accuracy at 88.95%.

Table 3. Classification of Saccade Features during WMC.

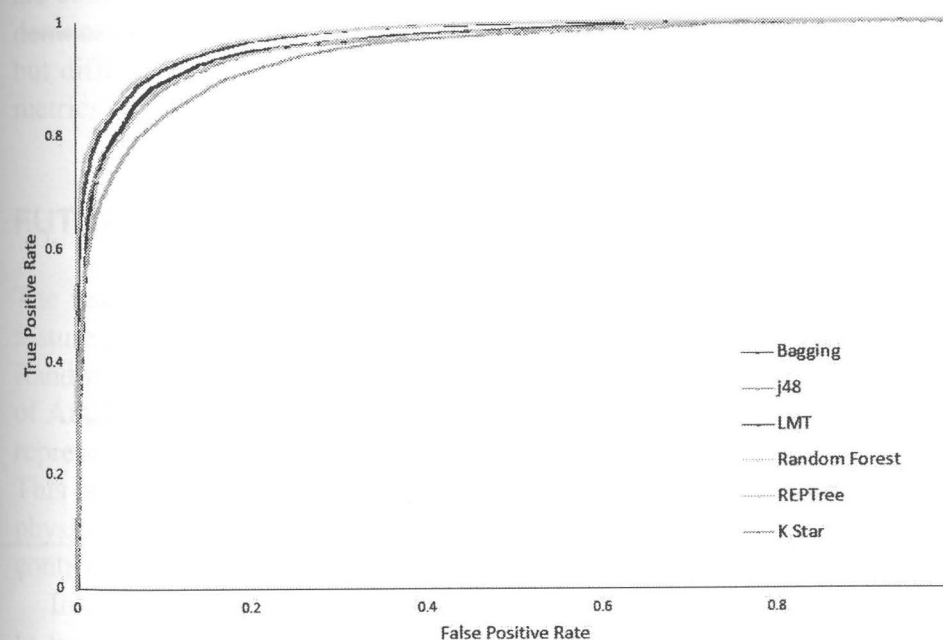
Classifier	Precision	Recall	F1	Accuracy
J48	0.89	0.89	0.89	88.95
LMT	0.89	0.89	0.89	89.51
RandomForest	0.91	0.91	0.91	91.14
REPTree	0.89	0.89	0.89	89.16
K*	0.86	0.86	0.86	85.98
Bagging	0.91	0.91	0.91	90.82

We generated a ROC graph (see Figure 3) for the classifiers we selected for the saccade feature set to investigate the performance metrics. The ROC curve shows that Random Forest classifier has the largest Area Under the Curve (AUC) meaning that it has the lowest error. The AUC of Random Forest classifier is 0.9114. Therefore, it has 91.14% chance of correctly distinguishing between ADHD and Non-ADHD. Random Forest is our top performing classifier offering the best trade-off in terms of the cost and the benefits for the saccade feature set.

Finally, table 4 provides results of the six top performing classifiers for the combination of saccade and fixations feature set. The Random Forest classifier yielded the highest percent accuracy at 91.11% indicating that the combination of fixation and saccade features collected during a RSPAN task classified a diagnosis of ADHD with greater than 90% accuracy. The K Star classifier yielded the lowest percent accuracy at 77.21%.

We generated a ROC graph (see Figure 4) to investigate the performance metrics for the classifiers we selected for the combination of saccade and

Figure 3. ROC Graph of the Top Performing Classifiers for Saccade Feature Set.



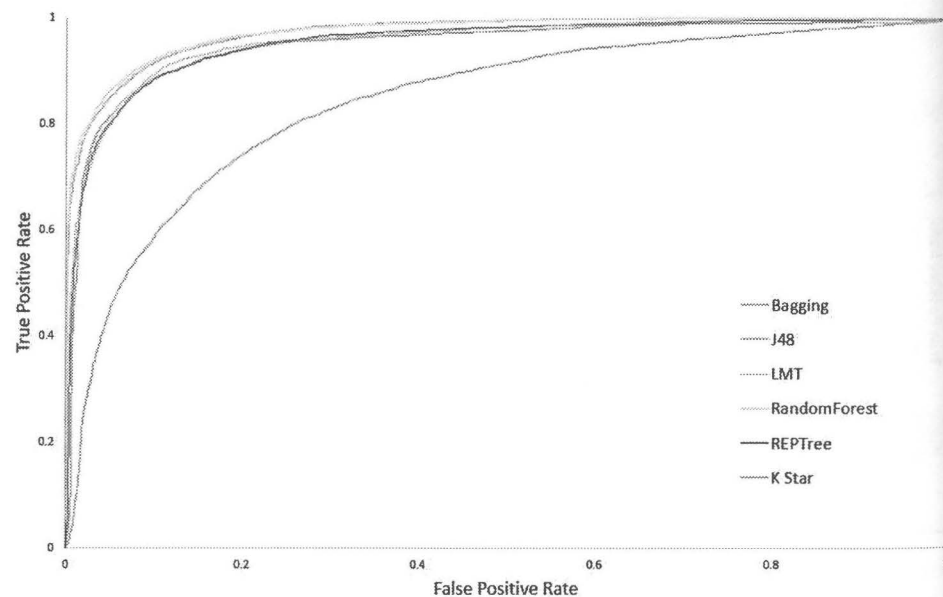
*For a more accurate representation see the electronic version.

Table 4. Classification of Eye Fixation and Saccade Features during WMC.

Classifier	Precision	Recall	F1	Accuracy
J48	0.89	0.89	0.89	89.19
LMT	0.89	0.89	0.89	89.91
RandomForest	0.91	0.91	0.91	91.11
REPTree	0.89	0.89	0.89	89.16
K*	0.77	0.77	0.77	77.21
Bagging	0.91	0.91	0.91	90.83

fixations feature set as well. The graph shows that even for the combination of saccade and fixations feature set, Random Forest is the top most performing classifier offering the best trade-off in terms of the cost and the benefits. The AUC of Random Forest classifier is 0.9111 meaning that it has 91.11% chance of correctly distinguishing between ADHD and Non-ADHD.

Figure 4. ROC Graph of the Top Performing Classifiers for the Combination of Fixation and Saccade Features Set.



*For a more accurate representation see the electronic version.

CONCLUSION

The purpose of this feasibility study was to determine if patterns of saccades and stability of fixations generated when completing a measure of WMC, the R-SPAN task, would create a feature set which could be used to differentiate a diagnosis of ADHD for adults. We used accuracy, precision, recall, and f-measure as the evaluation metrics. While fixation features, saccade features, and a combination of saccade and fixation features accurately predicted the classification of ADHD with an accuracy of greater than 78%, saccade features were the best predictors with an accuracy of 91%. These results are consistent with previous studies confirming significant differences in saccadic behaviors for people with ADHD (Lee et al. 2015; Roberts et al. 2011). During the R-SPAN task, the rapid movement of the eye across the scan path from one fixation point to the other yields a more accurate classification of ADHD than the ability to sustain gaze. These preliminary results indicate that detailed and discrete eye gaze metrics during a measure of attention control (i.e. WMC) provide unique indices of ADHD and offer physiological insight regarding cognitive resources underlying WMC, an important cognitive construct responsible for behavioral inhibition and attention monitoring. Moreover, they are consistent with previous investigations finding that adults with ADHD demonstrate similar broad visual attention patterns as adults without ADHD but different scan patterns (Krejtz et al. 2015) and different pupillometry metrics as a function of visual cue type (Michalek and Roche 2017).

FUTURE DIRECTIONS

The results of this feasibility study confirm the utility of eye movement feature set indexing WMC as a predictor of a diagnosis of ADHD in adults. RandomForest classifiers performed best in-terms of predicting a classification of ADHD with 91.14% percent accuracy by combining saccade feature set representing a physiological measure of visual attention during a WMC task. This project is a necessary first step in delineating a feature set of objective physiological biomarkers for important diagnostic criteria, including attention control, in adult ADHD.

In the future, we are interested in expanding the experimental studies to further analyze a fixation as well as saccade features set according to stimulus areas of interest using a larger sample size. Specifically, identifying

relationships between sentence units which determine sentence accuracy, the visual point of decision, the item to be remembered, and performance on the WMC measure. Using these areas of interest will generate a more detailed understanding of the relationships between covert measures of visual attention utilized during WMC tasks which also delineate a diagnosis of ADHD in adults.

REFERENCES

- Abibullaev, B., & An, J. (2012). Decision support algorithm for diagnosis of adhd using electroencephalograms. *Journal of Medical Systems*, 36(4), 2675–2688. doi:10.1007/10916-011-9742-x PMID:21671069
- Alderson, R. M., Kasper, L. J., Hudec, K. L., & Patros, C. H. (2013). Attention-deficit/hyperactivity disorder (ADHD) and working memory in adults: a meta-analytic review. *Neuropsychology*, 27(3), 287.
- Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of adhd. *Psychological Bulletin*, 121(1), 65.
- Biederman, J., Hammerness, P., Sadeh, B., Peremen, Z., Amit, A., Or-Ly, H., ... Faraone, S. (2017). Diagnostic utility of brain activity flow patterns analysis in attention deficit hyperactivity disorder. *Psychological Medicine*, 47(7), 1259–1270. doi:10.1017/S0033291716003329 PMID:28065167
- Blazey, R. N., Patton, D. L., & Parks, P. A. (2003). *ADHD Detection By Eye Saccades*. US Patent 6,652,458.
- Colflesh, G. J., & Conway, A. R. (2007). Individual differences in working memory capacity and divided attention in dichotic listening. *Psychonomic Bulletin & Review*, 14(4), 699–703. doi:10.3758/BF03196824 PMID:17972736
- Conway, A., Jarrold, C., & Kane, M. (2008). *Variation in working memory*. Oxford University Press. doi:10.1093/acprof:oso/9780195168648.001.0001
- Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12(5), 769–786. doi:10.3758/BF03196772 PMID:16523997
- Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9(4), 561.
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, 11(1), 19–23. doi:10.1111/1467-8721.00160
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of Experimental Psychology: General*, 128(3), 309.
- Fields, S. A., Johnson, W. M., & Hassig, M. B. (2017). Adult adhd: Addressing a unique set of challenges. *The Journal of Family Practice*, 66(2), 68–74. PMID:28222452
- Fostick, L. (2017). The effect of attention-deficit/hyperactivity disorder and methylphenidate treatment on the adult auditory temporal order judgment threshold. *Journal of Speech, Language, and Hearing Research: JSLHR*, 60(7), 2124–2128. doi:10.1044/2017_JSLHR-H-16-0074 PMID:28672285
- Greenhill, L. L. (1998). Diagnosing attention-deficit/hyperactivity disorder in children. *The Journal of Clinical Psychiatry*, 59, 31–41. PMID:9680051
- Hart, H., Chantiluke, K., Cubillo, A. I., Smith, A. B., Simmons, A., Brammer, M. J., ... Rubia, K. (2014). Pattern classification of response inhibition in adhd: Toward the development of neurobiological markers for adhd. *Human Brain Mapping*, 35(7), 3083–3094. doi:10.1002/hbm.22386 PMID:24123508
- Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, 130(2), 169.
- Komogortsev, O., Holland, C., Jayarathna, S., & Karpov, A. (2013). 2d linear oculomotor plant mathematical model: Verification and biometric applications. *ACM Transactions on Applied Perception (TAP)*, 10(4), 27.
- Komogortsev, O. V., Jayarathna, S., Aragon, C. R., & Mahmoud, M. 2010. Biometric identification via an oculomotor plant mathematical model. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*. ACM. 10.1145/1743666.1743679

- Krejtz, K., Duchowski, A., Szmidt, T., Krejtz, I., González Perilli, F., Pires, A., Vilaro, A., & Villalobos, N. (2015). Gaze transition entropy. *ACM Transactions on Applied Perception (TAP)*, 13(1), 4.
- Krejtz, K., Duchowski, A. T., Niedzielska, A., Biele, C., & Krejtz, I. (2018). Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze. *PloS One*, 13(9).
- Lee, Y.-J., Lee, S., Chang, M., & Kwak, H.-W. (2015). Saccadic movement deficiencies in adults with adhd tendencies. *ADHD Attention Deficit and Hyperactivity Disorders*, 7(4), 271–280. doi:10.1007/12402-015-0174-1 PMID:25993912
- Leigh, R. J., & Zee, D. S. (2015). The neurology of eye movements. Oxford University Press.
- Leigh, R. J., & Kennard, C. (2004). Using saccades as a research tool in the clinical neurosciences. *Brain*, 127(3), 460–477. doi:10.1093/brain/awh035 PMID:14607787
- Manevitz, L., & Yousef, M. (2007). One-class document classification via neural networks. *Neurocomputing*, 70(7-9), 7–9, 1466–1481. doi:10.1016/j.neucom.2006.05.013
- Marcano, J. L. L., Bell, M. A., & Beex, A. L. (2016). Classification of adhd and non-adhd using ar models. In *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the. IEEE*. 10.1109/EMBC.2016.7590715
- Michalek, A. M., Watson, S. M., Ash, I., Ringleb, S., & Raymer, A. (2014). Effects of noise and audiovisual cues on speech processing in adults with and without adhd. *International Journal of Audiology*, 53(3), 145–152. doi:10.3109/14992027.2013.866282 PMID:24456181
- Michalek, A. P., & And Roche, J. (2017). Pupil dilation as a measure of attention in adhd: A pilot study. *America*, 116, 2395–2405.
- Mostofsky, S. H., Lasker, A., Cutting, L., Denckla, M., & Zee, D. (2001). Oculomotor abnormalities in attention deficit hyperactivity disorder a preliminary study. *Neurology*, 57(3), 423–430. doi:10.1212/WNL.57.3.423 PMID:11502907

- Mueller, A., Candrian, G., Kropotov, J. D., Ponomarev, V. A., & Baschera, G.-M. (2010). Classification of adhd patients on the basis of independent erp components using a machine learning system. In *Nonlinear biomedical physics (Vol. 4)*. BioMed Central, S1. doi:10.1186/1753-4631-4-S1-S1
- Munoz, D. P., Armstrong, I. T., Hampton, K. A., & Moore, K. D. (2003). Altered control of visual fixation and saccadic eye movements in attention-deficit hyperactivity disorder. *Journal of Neurophysiology*, 90(1), 503–514. doi:10.1152/jn.00192.2003 PMID:12672781
- National Academies Of Sciences. E., Medicine. (2016). Improving diagnosis in health care. National Academies Press.
- Peng, X., Lin, P., Zhang, T., & Wang, J. (2013). Extreme learning machine-based classification of adhd using brain structural mri data. *PloS One*, 8(11).
- Richmond, J., & Nelson, C. A. (2009). Relational memory during infancy: Evidence from eye tracking. *Developmental Science*, 12(4), 549–556. doi:10.1111/j.1467-7687.2009.00795.x PMID:19635082
- Roberts, W., Fillmore, M. T., & Milich, R. (2011). Separating automatic and intentional inhibitory mechanisms of attention in adults with attention-deficit/hyperactivity disorder. *Journal of Abnormal Psychology*, 120(1), 223.
- Rommelse, N., Van Der Stigchel, S., Witlox, J., Geldof, C., Deijen, J.-B., Theeuwes, J., ... Sergeant, J. (2008). Deficits in visuo-spatial working memory, inhibition and oculomotor control in boys with adhd and their non-affected brothers. *Journal of Neural Transmission (Vienna, Austria)*, 115(2), 249–260. doi:10.1007/00702-007-0865-7 PMID:18253811
- Rönnerberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., ... Rudner, M. (2013). The ease of language understanding (elu) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience*, 7, 31. doi:10.3389/fnsys.2013.00031 PMID:23874273
- Rubia, K. (2018). Cognitive neuroscience of attention deficit hyperactivity disorder (adhd) and its clinical translation. *Frontiers in Human Neuroscience*, 12, 100. doi:10.3389/fnhum.2018.00100 PMID:29651240
- Tseng, P.-H., Cameron, I. G., Pari, G., Reynolds, J. N., Munoz, D. P., & Itti, L. (2013). High-throughput classification of clinical populations from natural viewing eye movements. *Journal of Neurology*, 260(1), 275–284. doi:10.1007/00415-012-6631-2 PMID:22926163

Van Der Stigchel, S., Rommelse, N., Deijen, J., Geldof, C., Witlox, J., Oosterlaan, J., ... Theeuwes, J. (2007). Oculomotor capture in adhd. *Cognitive Neuropsychology*, 24(5), 535–549. doi:10.1080/02643290701523546 PMID:18416506

Willcutt, E. G. (2012). The prevalence of dsm-iv attention-deficit/hyperactivity disorder: A meta-analytic review. *Neurotherapeutics; the Journal of the American Society for Experimental NeuroTherapeutics*, 9(3), 490–499. doi:10.1007/13311-012-0135-8 PMID:22976615

Xu, G., Strathearn, L., Liu, B., Yang, B., & Bao, W. (2018). Twenty-year trends in diagnosed attentiondeficit/hyperactivity disorder among us children and adolescents, 1997-2016. *JAMA Network Open*, 1(4), e181471–e181471. doi:10.1001/jamanetworkopen.2018.1471 PMID:30646132