

Multinomial Processing Models in Visual Cognitive Effort Diagnostics

Joshua D. Elkins

IUPUI

420 University Blvd.

joelkins@iupui.edu

Gahangir Hossain

IUPUI

420 University Blvd.

ghossain@iupui.edu

Abstract

The pupillary response has been used to measure mental workload because of its sensitivity to stimuli and high resolution. The goal of this study was to diagnose the cognitive effort involved with a task that was presented visually. A multinomial processing tree (MPT) was used as an analytical tool in order to disentangle and predict separate cognitive processes, with the resulting output being a change in pupil diameter. This model was fitted to previous test data related to the pupillary response when presented a mental multiplication task. An MPT model describes observed response frequencies from a set of response categories. The parameter values of an MPT model are the probabilities of moving from latent state to the next. An EM algorithm was used to estimate the parameter values based on the response frequency of each category. This results in a parsimonious, causal model that facilitates in the understanding the pupillary response to cognitive load. This model eventually could be instrumental in bridging the gap between human vision and computer vision.

1. Introduction

A critical goal of cognitive neuroscience and psychology is to determine connections between human thought processes and the resulting behavioral patterns. Cognitive load can be generally defined as the mental processing load or workload exerted during the performance of a cognitive task. An increase in cognitive load causes changes in autonomic physiological responses, such as electro-dermal activity, changes in heart rate, and changes in pupil diameter.

Changes in pupil diameter respond to cognitive load quickly and can reflect small differences in the load. Therefore, the pupillary response has been found to be an optimal secondary physiological measure of mental workload, due to its sensitivity to stimuli and high resolution. Without loss of generality, pupil dilation magnitude increases as the difficulty of a cognitive task increases [1]. The goal of this study was to diagnose the

cognitive effort involved with a task that was presented visually. Before discussing the methods used for this study, it is necessary to examine the instrumentation used, the data collected, and the neuroanatomical context of the model.

1.1. Instrumentation

Most modern pupillometry studies use remote eye trackers in order to measure changes in pupil diameter. There are two major types of eye trackers used: head-fixed and remote. Head-fixed camera eye trackers place the camera close to the subject's eye. In order to accomplish this, either the camera is head-mounted or the camera is desk-mounted and an apparatus such as a chin rest is used to restrain the subject's head. Remote eye trackers do not require the head to be fixed and are typically mounted on a desktop below eye level, yet still within the field of vision.

Both types of devices have their own advantages and disadvantages. Although head-fixed eye trackers have a relatively high accuracy, they may cause discomfort for the subject, resulting in artifacts. While remote eye trackers have a lower accuracy, they do not come into contact with the subject. Since remote eye trackers have a comparable precision, they are still widely used for the measurement of relative pupil dilations [1][2].

1.2. Collected Data

The pupillary data was adopted from Dr. Jeff Klingner's dissertation. During his study, he collected pupillary data from 12 students studying either Computer Science or Communications at Stanford University. 431 trials were performed, with number of trials per subject ranging from 32 to 43 [3].

During a 2 second accommodation period, each subject fixated on a target presented at the center of a computer screen. After the accommodation period, a mental multiplication problem was visually presented on the screen. The task had three levels of difficulty: easy, medium, and hard. For example 3×6 would be considered an easy task, while 15×19 would be considered a more difficult task. The subjects were given 5 seconds to respond. They typed their answers on a low-contrast, on-

screen keypad. The mental multiplication problems were randomly selected and were categorized as either easy, medium, or hard. Klingner maintained a constant luminance for the graphical user interface and the entire room. The pupillary response was recorded with the Tobii T120 remote desktop eye tracker. The data was sampled at a rate of 50 Hz [3]. The average pupillary response across all trials, for each level of difficulty is shown below in Figure 1.

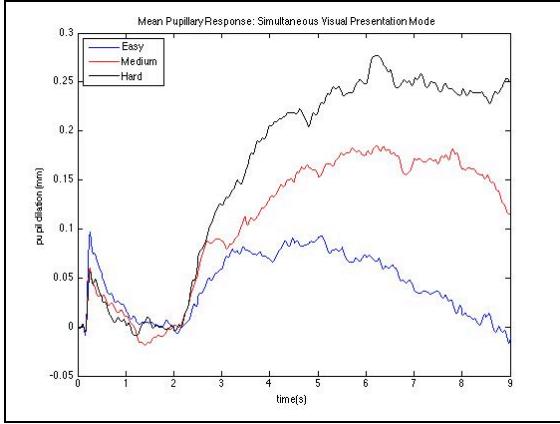


Figure 1: The average pupillary response across all trials, for each level of difficulty (easy, medium, and hard) is shown above. This figure was recreated using raw data collected by Dr. Jeff Klingner. Signal processing (described in more detail in the methods) was performed in MATLAB [3].

1.3. Neuroanatomical Context

An understanding of neural anatomy is necessary in order to determine the relationship between changes in cognitive load and the pupillary response. This section proposes a possible neurological pathway that is involved with the pupillary response to cognitive load.

A visual stimulus enters through the pupil and is projected onto the retina. Photoreceptors within the retina convert light energy into electrical energy in a process known as photo-transduction. The electrical signal propagates along the optic nerve and crosses over at the optic chiasm. The signal travels through the thalamic Lateral Geniculate Nucleus (LGN) to the Primary Visual Cortex. From the Primary Visual Cortex, the signal is transferred to the prefrontal lobe where the multiplication operation can be performed. During this process the electrical signal travels to the Hippocampus. From the Hippocampus, the signal travels along the Para hippocampal gyrus to the central nucleus of the Amygdala. The signal reaches the Locus Ceruleus (LC) (located in the pons) via efferent fibers of the Amygdala. The LC releases norepinephrine, which eventually reaches the neuromuscular junction of the

dilator pupillae and the ciliary ganglion. The dilator pupillae muscles have norepinephrine receptors that, when activated, cause the pupil to dilate [4][5]. The neuroanatomical context of the pupillary response to cognitive load is illustrated below in Figure 2.

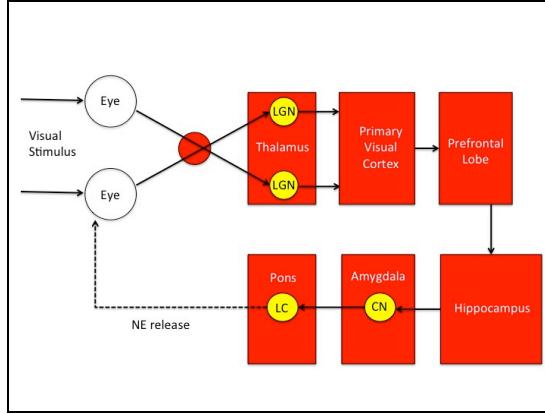


Figure 2: A diagram illustrating a possible pathway involved in the pupillary response to cognitive load is shown above.

An understanding of the neurological pathways involved with the pupillary response to cognitive load facilitates the construction of a visual cognitive diagnostic model. Computer vision and pattern recognition algorithms use black, white, or grey box methods of evaluation. A black box system has known inputs and outputs, but the internal structure is either not well known or completely unknown. In contrast the internals of a white box system are completely known. In a grey box system, there is some insight of the internals of the system, yet there are still unknown internal components [7].

A cognitive diagnostic model using the pupillary response can be thought of as a grey box system. The inputs are the presented mental multiplication tasks and the outputs are the subjects' responses to the problem. The internal components that are unknown are the neurological mechanisms involved. If the neural pathways involved in the cognitive pupillary response are understood, then the pupillary dynamics can provide insight into the system internals. A block diagram of a cognitive diagnostic model using the pupillary response is shown in Figure 3. Recently computer vision has been transitioning into a more robust cognitive vision, allowing for the AI system to learn, adapt, and weigh alternative solutions. Machine learning algorithms, such as artificial neural networks, can learn from the parameters determined by the cognitive diagnostic model and allow the system to process visual information similar to how a human would.

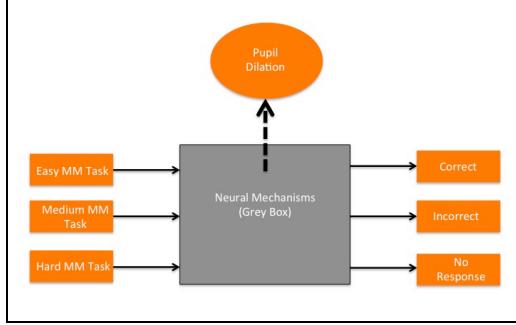


Figure 3: The block diagram above illustrates the visual cognitive diagnostic model as a grey box system. The inputs are the presented mental multiplication tasks of varying difficulty. The outputs are whether the subject answered correctly, incorrectly, or provided no response. The measured changes in pupil diameter provide insight into the internal components of the system.

2. Methods

The following section provides a description of methods used for this study. Though listed sequentially, some of the steps involved in data analysis were performed sequentially.

2.1. Signal Processing of the Raw Pupillary Data

Signal processing of the raw pupillary data was performed in MATLAB. The methods are similar to those performed by Klingner et al. The raw pupillary data consisted of the pupil diameter measured from both the left and right eye of each subject. All blinks were removed from each trial. A linear interpolation was performed at each resulting gap.

In order to convert the absolute pupil diameter data to relative pupil dilation data, a baseline subtraction was performed. To accomplish this, the pupil diameter was averaged over the last 20 samples of the pre-stimulus accommodation period. This average was considered the baseline pupil diameter. To obtain the relative pupil dilation signal, the baseline pupil diameter was subtracted from the absolute pupil diameter signal.

There is an abundant amount of high frequency, instrumentation noise involved with a remote eye tracker when measuring the pupillary response. This can be due to eye drift, tremors, and a non-spherical eye shape. The pupillary response is a low frequency signal. Therefore, a 10 Hz, low-pass FIR filter was used in order to reduce the high frequency noise of the signal. After filtering the signal, the left and right eyes were averaged. The signal processing involved is illustrated in Figure 4.

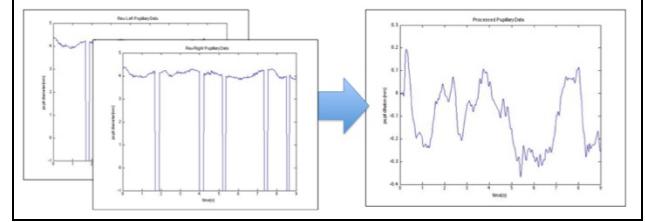


Figure 4: The diagram above illustrates the signal processing involved in this study.

2.2. Creation of Bins

All pupillary data trials for each level of difficulty were divided into three bins: small, medium, and large. There were three types of divisions used, which are listed below in Table 1.

Division Type #	Description
1	Mean \pm Standard Deviation (SD)
2	Mean \pm Standard Error (SEM)
3	Mean \pm Midpoint Between SD and SEM

Table 1: A list of the three types of bin divisions is shown in the above table.

A method used to evaluate each bin division type will be discussed later in this section. Figures 5 and 6 illustrate bin division types 1 and 2, respectively, for all trials involving the presentation of an easy multiplication task.

For each bin, the data was further sorted into 3 response categories: correct, incorrect, and no response. As a result the entire data was sorted into a total of 27 unique response categories. The response frequencies were determined for each category. Data sorting was performed in MATLAB.

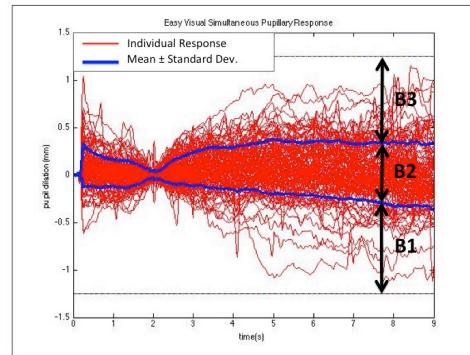


Figure 5: Division type 1 for easy pupillary data is illustrated above.

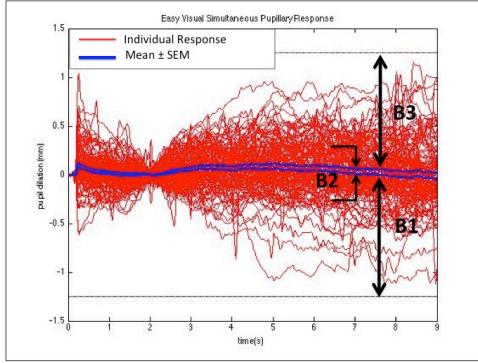


Figure 6: Division Type 2 for easy pupillary data is illustrated above.

2.3. Multinomial Processing Trees

A major challenge in developing a cognitive model is being able to quantify latent cognitive states. A Multinomial Processing Tree (MPT) was used as an analytical tool in order to disentangle and predict separate cognitive processes, with the resulting output being a change in pupil diameter. An MPT model describes observed response frequencies from a set of response categories. The “root” of the tree is presented stimulus. Each “branch” of the tree represents an estimated parameter value. The parameter values of the MPT model are the probabilities of moving from one latent state to the next. The “leaves” of the MPT are the known response frequencies. While ad hoc analyses, such as ANOVA, are limited to only testing hypotheses on observed data, MPT models decompose the data into latent cognitive processes [8].

multiTree, a computer program developed at the University of Mannheim, was used to create the MPT models used for this study. The software uses an Expectation Maximization (EM) algorithm in order to estimate the parameter values of any given MPT model. As the name implies, this algorithm can be split into an expectation phase and a maximization phase. Before the EM parameter is implemented, it is necessary to initialize the parameter values. The software automatically initializes all parameter values at 0.5. During the expectation phase, the parameter vector of the previous trial is used to estimate the expected frequency values. During the maximization phase, the parameter values are estimated based on the expected frequencies from the

expectation phase. The algorithm iterates until convergence.

There are a large variety of criteria that can be used for the final model selection. Each one has its strengths and weaknesses. Therefore, multiple criteria should be observed to determine the “best” final model.

G^2 is a goodness-of-fit measure that is the discrepancy between the statistical predictions and the observed data. This is only a measure of goodness of fit. It does not take flexibility of the model into account.

Akaike’s Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are measures that do take flexibility of the model into account by adding punishment factors to the original G^2 value. The drawback to these measures is that they ignore the differences between various models’ functional form. The equations for these two criteria are shown below:

$$AIC = G^2 + 2S \quad (1)$$

$$BIC = G^2 + \ln(N) * S \quad (2)$$

The Fisher Information Approximation (FIA) is a measure that favors simpler models. It takes into consideration the goodness-of-fit, flexibility, and the model’s functional form. The criterion is calculated using the following formula:

$$FIA = \frac{1}{2} * G^2 + \frac{S}{2} \ln\left(\frac{N}{2}\right) + \int \sqrt{\det(I(\Theta))} d\Theta \quad (3)$$

As shown above, this is a computationally intensive calculation that requires software [8][9]. It is desirable to minimize FIA, AIC, and BIC.

This study can be divided into two phases: the best bin division phase and the model-fitting phase. During the best bin division phase, the simplest possible MPT model was generated using the response frequencies of each of the three bin division types. The model selection criteria were compared to determine the best bin division.

During the model-fitting phase, a more complex model that provides a greater amount of cognitive information was fitted to the best bin division data. The multinomial processing trees used for the best bin division phase and the model-fitting phase are shown in Figures 7 and 8, respectively.

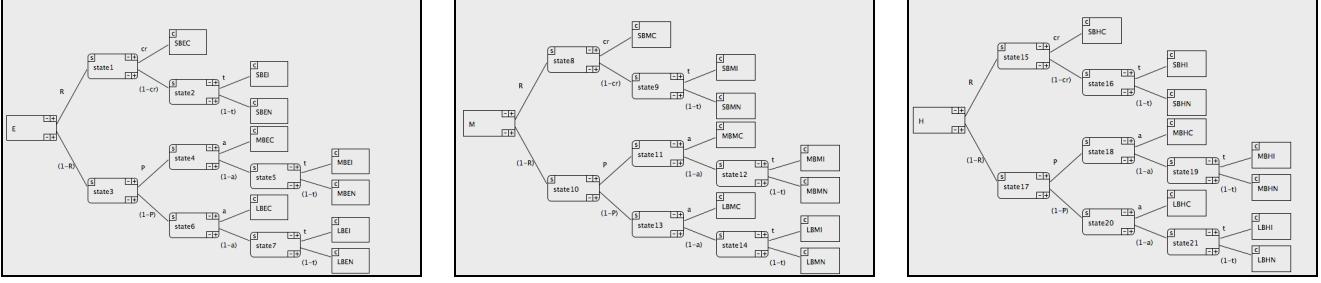


Figure 7: The multinomial processing tree model for the best bin division phase is shown above. There were three trees: Easy (left), Medium (center), and Hard (right).

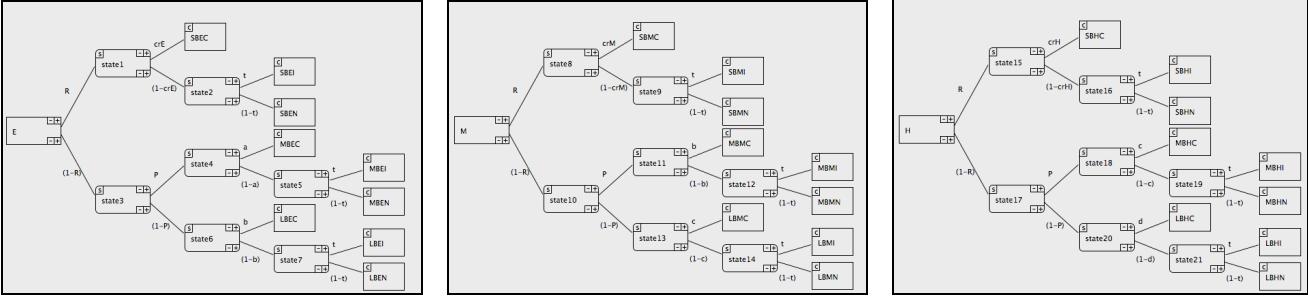


Figure 8: The multinomial processing tree model for the model-fitting phase is shown above. There were three trees: Easy (left), Medium (center), and Hard (right).

The description of each parameter of the MPT model generated during the best bin division phase is shown below in Table 2.

Parameter	Description
R	This is the probability that the subject will recognize the mental multiplication problem, recalling it from long-term memory.
P	This is the probability that the subject will perceive that the problem is at the level of difficulty.
a	This is the probability that the subject knows how to solve the problem.
cr	This is the probability that the subject can correctly recall the answer a problem that he or she recognizes.
t	This is the probability that the subject is able to meet the time constraints of the task.

Table 2: A list describing each parameter of the MPT generated during the best bin division phase is shown in the above table.

The description of each parameter of the MPT model generated during the model-fitting phase is shown below in Table 3.

Parameter	Description
R	Same as Best Bin Division Phase
P	Same as Best Bin Division Phase
a	Solving Probability when Perceived as Easy Difficulty
b	Solving Probability when Perceived as Medium Difficulty
c	Solving Probability when Perceived as Hard Difficulty
d	Solving Probability when Perceived as Extremely Difficult
crE	Correct Recall Probability for Easy
crM	Correct Recall Probability for Medium
crH	Correct Recall Probability for Hard
t	Same as Best Bin Division Phase

Table 3: A list describing each parameter of the MPT generated during the model-fitting phase is shown in the above table.

3. Results

The following section provides the results of both the best bin division phase and the model-fitting phase.

3.1. Best Bin Division Phase

Table 4 compares the model selection criteria for each type of bin division.

Bin Division Type #	AIC	BIC	FIA
1	1110	1129	563.7
2	1228	1247	622.9
3	1325	1344	671.4

Table 4: The above table compares the model selection criteria for each type of bin division.

3.2. Model-Fitting Phase

The parameter values of the model-fitting phase are shown in Table 5. The table also includes SEM values and 95% confidence limits.

Parameter	Mean Value	SEM	95% Lower CL	95% Upper CL
R	0.147	0.018	0.111	0.182
P	0.840	0.020	0.801	0.881
a	0.979	0.015	0.950	1.000
b	0.857	0.032	0.794	0.920
c	0.446	0.052	0.344	0.547
d	0.600	0.126	0.352	0.848
crE	0.941	0.057	0.829	1.000
crM	0.917	0.056	0.806	1.000
crH	0.571	0.132	0.312	0.830
t	0.682	0.051	0.583	0.781

Table 5: The above table compares the parameter values during the model-fitting phase.

All parameter values of interest with their SEM bars are illustrated in a bar graph shown in Figure 9. All parameter values with their confidence intervals are illustrated in a bar graph shown in Figure 10. Note that the parameters are sampled from a Beta distribution. Therefore a t-test or a standard ANOVA test is not applicable, since they both have a normality assumption.

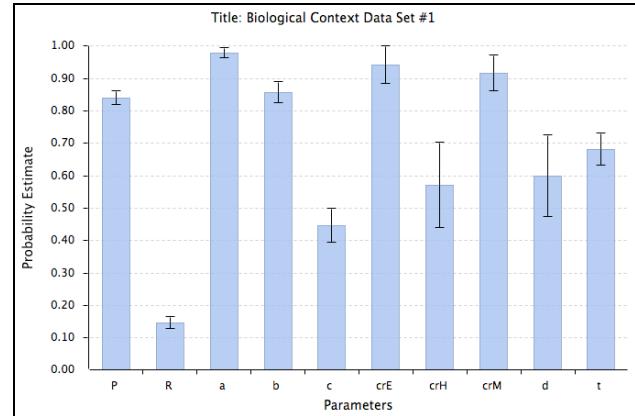


Figure 9: A bar graph of the parameter values with their SEM bars is shown above.

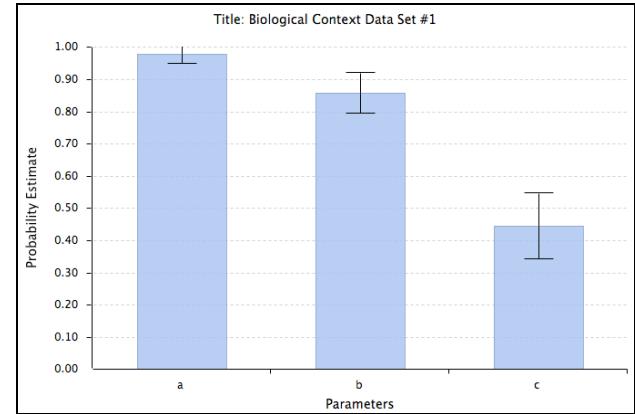


Figure 10: A bar graph of the parameter values with their confidence intervals is shown above.

4. Discussion

Based on the model selection criteria, the best bin division type was division type 1, which uses the standard deviation of the pupillary response to separate the bins. The response frequencies for the model-fitting phase were based off bin division type 1.

Parameters a, b, and c were found to be significantly different from one another. Both their SEM bars and confidence intervals did not overlap, suggesting statistical significance. Based on the bar graph of these parameters of interest shown in Figure 10, the probability that the subject knows how to solve the mental multiplication task decreases with the level of perceived difficulty. Statistical significance was determined by using the SEM and the confidence intervals of the parameter values. It should be noted that a more formal and robust significance test for a Beta distribution should be implemented in future studies.

The parameters used in the multinomial processing tree model have more of a psychological context. More research is necessary to determine a model that has a more neurobiological context.

Using a MPT model is a robust way of diagnosing visual cognitive effort. The model determined in this study was a parsimonious, causal model that facilitates in the understanding of the pupillary response to cognitive effort. Future research will be devoted to designing an MPT that further aligns with the neurological pathways involved, with the parameter values representing the probability of an action potential propagating to the next subcomponent of the visual pathway. Therefore, the use of an MPT as a diagnostic tool for visual cognitive effort could be instrumental in bridging the gap between human vision and computer vision.

References

- [1] Klingner, J., “Background”, in *Measuring Cognitive Load During Visual Tasks by Combining Pupillometry and Eye Tracking*, Stanford University, 2010, ch.2, pg. 7-15.
- [2] Klingner, J., “Remote Eye Tracker Performance”, in *Measuring Cognitive Load During Visual Tasks by Combining Pupillometry and Eye Tracking*, Stanford University, 2010, ch.3, pg. 17-28.
- [3] Klingner, J., “From Auditory to Visual”, in *Measuring Cognitive Load During Visual Tasks by Combining Pupillometry and Eye Tracking*, Stanford University, 2010, ch.5, pg. 41-61.
- [4] Haines, D., “The Visual System”, in *Fundamental Neuroscience for Basic and Clinical Applications*, Philadelphia, PA, 2013, ch.20, pg. 267-286.
- [5] Haines, D., “The Limbic System”, in *Fundamental Neuroscience for Basic and Clinical Applications*, Philadelphia, PA, 2013, ch.31, pg. 431-441.
- [6] Hubert, M., Vandervieren, E., “An Adjusted Box Plot for Skewed Distributions”, in *Section of Statistics*, Katholieke Universiteit Leuven, November 2006, pg. 1-10.
- [7] Cross Check Networks, “SOA Testing Techniques”, [Online], Available: http://www.crosschecknet.com/soa_testing_black_white_gray_box.php, 2015.
- [8] Singmann, H., “MPTinR: Analysis of Multinomial Processing Trees in R”, *Behavior Research Methods*, June 2013, 45, 2, pg. 560-575.
- [9] Moshagen, M., “multiTree: A Computer Program for The Analysis of Multinomial Processing Tree Models”, *Behavior Research Methods*, 2010, 42, 1, pg. 42-54.
- [10] Hossain G., Yeasin M., “Understanding Effects of Cognitive Load from Pupillary Response Using Hilbert Analytic Phase”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2014, pg. 375 – 380.
- [11] Klingner J., Tversky B., Hanrahan P., “Effects of visual and verbal presentation on cognitive load in vigilance, memory,

and arithmetic tasks”, *Psychophysiology*, March 2011, 48, 3, pg. 323 – 336.