Online Journal of Communication and Media Technologies, 2024, **14**(1), e202413
e-ISSN: 1986-3497

OPEN ACCESS

Research Article

Cognitive effort assessment through pupillary responses: Insights from multinomial processing tree modeling and neural interconnections

Gahangir Hossain [1*]
0000-0002-8205-4939

Joshua D. Elkins [2]
0000-0003-0191-5512

[1] University of North Texas, Denton, TX, USA
[2] Indiana University–Purdue University Indianapolis, Indianapolis, IN, USA
[*] Corresponding author: gahangir.hossain@unt.edu

Citation: Hossain, G., & Elkins, J. D. (2024). Cognitive effort assessment through pupillary responses: Insights from multinomial processing tree modeling and neural interconnections. *Online Journal of Communication and Media Technologies, 14*(1), e202413. https://doi.org/10.30935/ojcmt/14196

ARTICLE INFO

Received: 10 Aug 2023

Accepted: 17 Jan 2024

ABSTRACT

The pupillary responses of humans exhibit variations in size, which are mediated by optic and oculomotor cranial nerves. Due to their sensitivity and high resolution of pupillary responses, they are used for a long time as measurement metrics of cognitive effort. Investigating the extent of cognitive effort required during tasks of varying difficulty is crucial for understanding the neural interconnections underlying these pupillary responses. This study aims to assess human cognitive efforts involved in visually presented cognitive tasks using the multinomial processing tree (MPT) model, an analytical tool that disentangles and predicts distinct cognitive processes, resulting in changes in pupil diameter. To achieve this, a pupillary response dataset was collected during mental multiplication (MM) tasks and visual stimuli presentations as cognitive tasks. MPT model describes observed response frequencies across various response categories and determines the transition probabilities from one latent state to the next. The expectation maximization (EM) algorithm is employed with MPT model to estimate parameter values based on response frequency within each category. Both group-level and individual subject-to-subject comparisons are conducted to estimate cognitive effort. The results reveal that in the group comparison and with respect to task difficulty level, that subject's knowledge on MM task influences the successfully solve the problem. Regarding individual analysis, no significant differences are observed in parameters related to correct recall, problem-solving ability, and time constraint compliance. However, some significant differences are found in parameters associated with the perceived difficulty level and ability to recall the correct answers. MPT model combined with EM algorithm constitutes a probabilistic model that enhances pupillary responses identification related to the cognitive effort. Potential applications of this model include disease diagnostics based on parameter values and identification of neural pathways that are involved in the pupillary response and subject's cognitive effort. Furthermore, efforts are underway to connect this psychological model with an artificial neural network.

Keywords: cognitive effort, mental multiplication, multinomial processing trees, pupillary dynamics

INTRODUCTION

Cognitive effort plays a significant role in everyday life, influencing cognitive task performance and having implications for healthy and disordered functioning across a wide range of tasks, including arithmetic and

This paper is a journal version of the workshop paper "Multinomial processing models in visual cognitive effort diagnostics" by Joshua D. Elkins & Gahangir Hossain at CVPR Workshops 2015 (pp. 9-15).

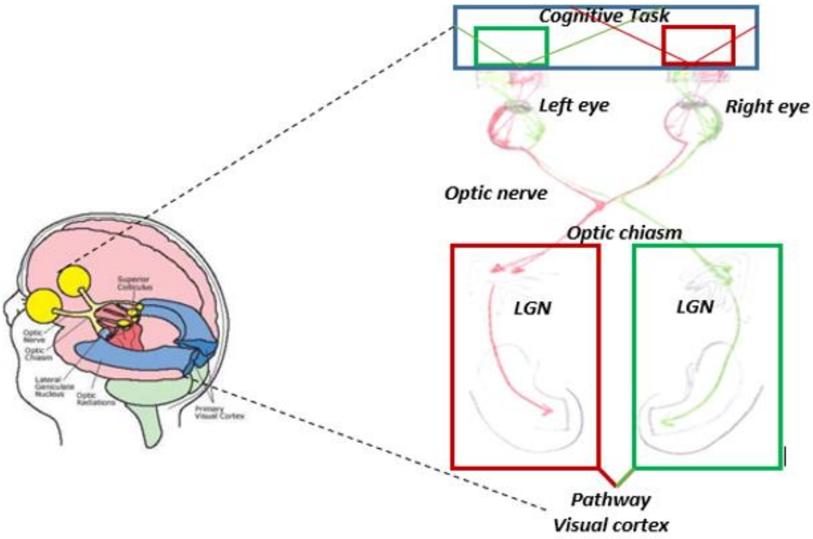

Figure 1. From pupillary response to mental workload (cognitive load) neural pathway (pathway [left] & amplified visual processing [right]) (Source: Authors)

decision-making. Despite its importance, computational modeling of cognitive effort diagnosis and understanding the core mechanisms underlying human experience, decision-making, and cognitive performance have been limited. One contributing factor to this limitation is the lack of a clear, real-time, and operational diagnosis of cognitive effort. Therefore, the objective of this paper is to adopt a computational source localizing model for cognitive diagnosis, focusing on effort-based decision-making.

Advanced psychology and cognitive neuroscience incorporate the connections between human thought processes and resulting behavior, for example, various eye activities including pupil dilation, eye gaze, eye wandering, heart rate variability, sweating rate, and more. The amount of mental processing needed in processing a cognitive task is known as cognitive effort, which is essential in everyday human activities. Cognitive effort depends on various aspects of task complexities or difficulties, this results in cognitive load. This change is represented through various biological responses, such as heart rate fluctuations, electroencephalograms (EEG), and alterations in pupil diameter (Klingner, 2010). The pupillary response, reflecting changes in cognitive effort rapidly and with sensitivity, has been identified as an effective measure of cognitive workload. Specifically, the cognitive task difficulty increases the magnitude of pupil dilation (Klingner, 2010; Sensi et al., 1999).

Understanding neural anatomy is crucial to establishing relationships among cognitive effort, pupillary responses, and changes in cognitive workload. Analyzing neurological pathways (visual or auditory) involved in human cognitive processing and their connection to response behaviors, such as the pupillary response, is essential in this context. The process begins with a visual signal entering the center eye area like the pupil and extended to the inner side as in the retina. Through the photo transition process, the light energy is converted into electrical stimuli within retina's photoreceptors. These electrical signals propagate through the optic nerve and cross over at the optic chiasm. Afterwards, they reach the primary visual cortex (PVC), traveling through the thalamic lateral geniculate nucleus (Haines, 2013).

In a highly engaged cognitive task, like the mental multiplication (MM) task, information is entered through human eye transfers from PVC to the prefrontal lobe. The operation of MM task takes place in frontal lobe. Internally, electrical signals travel to another part of the eye known as hippocampus. Later these electrical signals spread towards the para hippocampal gyrus, and central nucleus of the amygdala. The signal reaches another part of the internal eye known as the pons or locus coeruleus (LC), through amygdala's fibers. This is the mechanism of pupil dilation occurrence–by LC releasing norepinephrine and reaching the neuromuscular junction (Haines, 2013), the process is shown in **Figure 1** (Haines, 2013; Privitera et al., 2010). The performance in mental multiplication tasks is contingent upon the level of task difficulty with scaling (Griffith & Kalish, 2021), fluctuation of pupil (Ohtsuka et al., 1988; Sensi et al., 1999), pupil dilation on visual perception (Privitera et al., 2010).

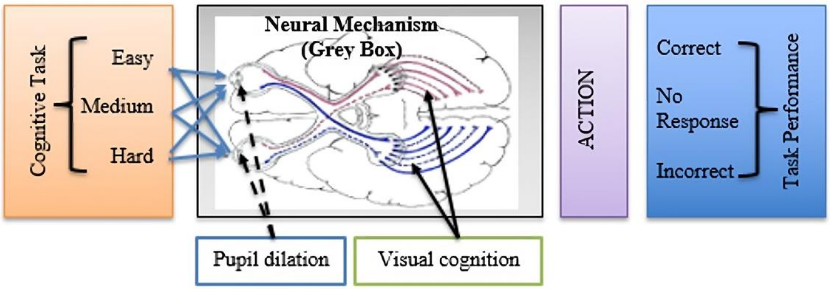

Figure 2. A grey box modeling of visual cognition diagnosis (Source: Authors)

Visual Cognitive Diagnostic Model

A visual cognitive diagnostic (VCD) model is the neural pathway, which is responsible for generating pupillary responses with respect to mental workload or cognitive load. This model offers a foundation for developing visual cognition. In connection with this VCD model, computing world utilizes three evaluation methods: black box, white box, or grey box methods (Cross Check Networks, 2015). In the context of cognitive diagnosis based on pupillary responses, it can be viewed as a grey box modeling approach. The inputs to this model can be cognitive tasks, and the outputs responses from the subject to these cognitive tasks and the problems. The unknown internal components refer to the involved neural mechanisms. Understanding the neural mechanism underlying the pupillary response with respect to cognitive task can shed light on the system's internal processes. **Figure 2** illustrates a framework of the cognitive diagnostic model incorporating their pupillary responses.

The visual cognitive diagnosis model is shown in **Figure 2**. MM tasks with varying difficulty are presented as the inputs and the subjects answers with feedback are presented as the output. The internal component of the system is reflected through the measured changes in pupil diameter (Cross Check Networks, 2015; de Gee et al., 2014; Einhäuser et al., 2008) except the planning disorderness (Köstering, 2012).

The paper is structured as follows: The introduction sets the stage by presenting the basic concept of human cognitive efforts, followed by an explanation of a visual cognitive modeling model in this section. Then, the paper delves into the current state of work concerning cognitive effort estimation, linking it to cognitive efforts and visual cognitions, particularly those stimulated visually (e.g., MM tasks). We then detail the research methodologies, encompassing datasets and identification of bins, while elaborating on the proposed multi-nominal processing tree-based effort estimation process. Next, we present processing outcomes, showcasing varied bins and visual cognitive efforts through pupillary dynamics. After that we engage in a discussion of implications arising from processed results. Finally, we conclude the paper, summarize findings, present a conceptual model, and outline potential future directions for research in this domain.

LITERATURE REVIEW

Bridging neurobiological and psychological analysis in the form of cognitive science plays a key role in cognitive effort bridge the gap between mental processing and resulting activity patterns. One crucial aspect in psychology is the cognitive load assessment, which generally refers to the processing of mental workload through assessing the limited capacity of mental resources versus the variation of cognitive task and their performance (Klingner, 2010). Cognitive responses, such as changes of signals in the autonomic nervous system, can serve as proxies for measuring cognitive load. These responses encompass EEG, eye activity, heart rates variability, and the pupil dilation and responses. Pupil diameter changes quickly in response to cognitive load and can reflect even subtle variations in the load. Research by Jeff Klingner compared the precision of a Tobii 1750 remote eye tracker to a reference neurotics VIP-200 pupillometer. Although Tobii 1750 had lower precision, it was still deemed acceptable for task-averaged cognitive pupillometry (Jepma & Nieuwenhuis, 2011). Pupil diameter measurements were found to be larger in exploration trials compared to exploitation trials, and the baseline pupil diameter increased with the level of exploration (Jepma & Nieuwenhuis, 2011). Spontaneous pupillary fluctuations in the time-frequency domain were also analyzed, considering factors like respiration and changes in human blood pressure. The authors utilized the short-time Fourier transform to compute power spectral density function of the pupillary response (Nowak et al., 2008).

Besides these literatures, Bishara and Payne (2008) employed a tree-based model named the multinomial processing tree (MPT) model to study weapon misidentification based on race. Another tree-based model, known as the process dissociation model, was introduced, which is determined to be the optimal model using the model selection criteria. Artificial neural networks (ANNs), inspired by biological neurons, process information with nodes representing cell bodies, inputs akin to dendrites, and outputs representing axons. The weighted inputs are summed at each node, and a nodal activation transfer function emulates the firing rate of a neural cell, leading to an output only if a certain threshold is reached (Stergiou & Siganos, 1996). Chen and Magdon-Ismail (2006) proposed a framework for learning to price complex options by utilizing MPTs.

Recently random effects MPT models with a maximum likelihood approach is proposed (Miles et al., 2023), which can be adopted in robust cognitive effort estimation. To investigate sleepiness and sleep quality and their relations MPT models are also investigated (Böhm et al., 2020), which is event-based. Nestler and Erdfelder (2023) sketch out several major concepts connected to sociology and cognition, in terms of response conflict and MPT.

The present work explores the application of MPT model (Singmann, 2010) in effective cognitive effort diagnosis for cognitive tasks with varying difficulty levels. It argues that MPT models and their variants offer intuitive cognitive effort diagnosis, supporting a neurocognitive-focused research strategy. The paper outlines cognitive tasks with varying difficulty levels and connects them to pupillary dynamics (pupil dilation) as an analysis opportunity Beatty (1982). It emphasizes the benefits of adopting MPT model with the expectation maximization (EM) algorithm for identifying markers of cognitive efforts (Hossain & Elkins, 2018). Additionally, the paper presents findings from parsing the pupillometry dataset into bins corresponding to standard deviations and suggests future research directions leveraging the potential of cognitive effort diagnosis with MPT variants (e.g., binary MPT, GPT model, etc.). The findings demonstrate the feasibility of human cognitive processing and the comprehensive application of MPT models and other cognitive processing methods within this domain. The paper discusses methodological and conceptual benefits of applying MPT models in cognitive effort estimation and modeling human cognitive load dynamics.

Recent studies, such as da Silva et al. (2021), focus on a task-switching paradigm to understand how individuals manage cognitive effort towards their goals. Their findings strongly indicate that task-evoked pupillary responses (TEPRs) correspond with increased effort during task switches, suggesting pupillometry as a promising marker of cognitive endeavor. However, accurately measuring cognitive effort remains a challenge. Reilly et al. (2019) suggest that TEPRs serve as a nuanced indicator for the intensity and cognitive demands across various mental processes, including problem-solving and memory retrieval. Their study reveals that TEPRs operate independently of baseline size, displaying a consistent linear scaling pattern across different tasks and lighting conditions. This study reevaluates past pupillometry methodologies and points toward future cognitive research methodologies. Additionally, Alsobeh and Shatnawi (2023) and Jarrah et al. (2021) propose an integrated model utilizing model checking and the BIP component model to enhance IoT security. This model incorporates BIP-based components for data security, analytics, threat detection, and continuous monitoring, aiming to fortify security in IoT systems against issues such as counterfeit data and malware threats. Notably, it also considers human factors like TEPRs in human-inspired home cognitive security.

METHODS

Both a group and individual subject-to-subject analyses were performed in this study. The following section explains the data used in this study to illustrate the visual cognitive effort through a cognitive task (MM task), cognitive signal processing, statistical analysis and MPT modeling.

Data

The pupillary data utilized in this study originated from research conducted by Klingner et al. (2011). The study involved 12 participants studying computer science or communications at Stanford University, and a

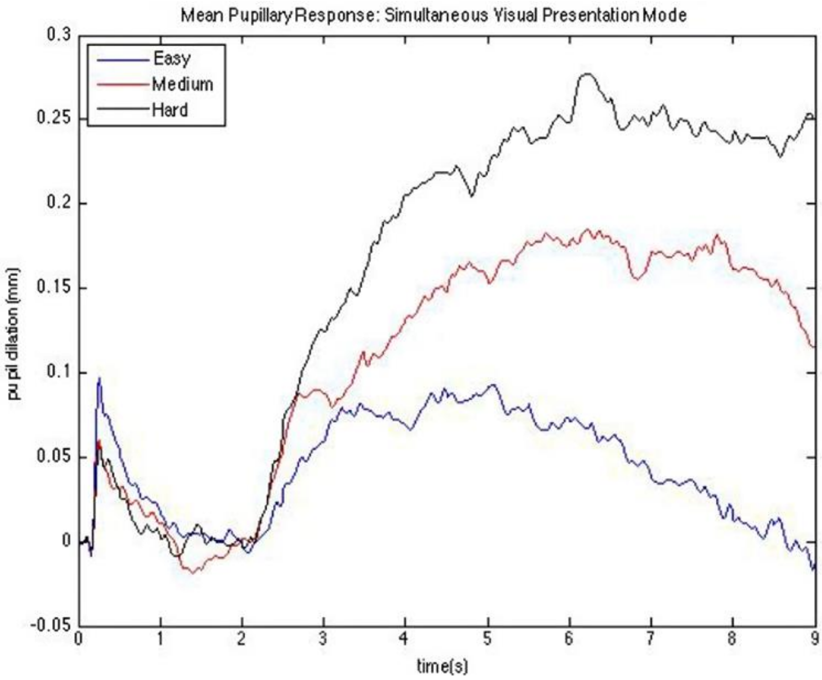

Figure 3. Average change of pupil diameters across all trials with varying difficulty levels (E, M, & H) (Klingner, 2010)

total of 431 trials were performed, with each subject completing between 32 to 43 trials (Klingner, 2010). Each subject underwent two seconds adjustment time while fixating on the visual cognitive task displayed at the focus point on the screen, in the center. Following this time, a visual cognitive task (MM) was presented on the screen. Each cognitive task allows the subject five seconds to response. The subject can answer or response using their keyboard or the on-screen keypad. Various cognitive tasks (MM problems) were selected randomly, which were categorized into three classes easy (E), medium (M), or the hard (H), based on their level of difficulties. According to Klingner et al. (2011), a consistent luminance was kept for both the user interface and the experimental room. The change in pupil size was recorded using Tobii T120 remote desktop eye tracker, with data being sampled at a rate of 50 Hz (Klingner, 2010). **Figure 3** displays the changes in pupil sizes with respect to trials for varying difficulty levels.

Signal Processing on Raw Pupillary Data

In signal processing MATLAB is used, followed from the similar methods used by Klingner et al. (2011). The initial raw data included the subject's left and right eye pupil diameter. Data quality was improved with blink removals and missing value filled with interpolation. Absolute pupil diameter data to relative pupil dilation data are converted with a baseline subtraction. The baseline pupil diameter was computed by averaging the pupil diameter over the last 20 samples of the pre-stimulus accommodation period. By subtracting this baseline value from the absolute pupil diameter signal, the relative pupil dilation signal was obtained.

Remote eye trackers data includes instrumentation noise in measuring the pupillary response, which could be affected by eye related outer activities including drift, tremors, and the non-spherical shape of the eye. The pupillary response itself is characterized by a low-frequency signal. Recognizing that the pupil diameter of the left and right eye was highly correlated at low frequencies, as observed by Klingner (2010), a 10 Hz low-pass finite impulse response (FIR) filter was applied to eliminate or reduce noise and improve the data quality.

This task was accomplished by using *fir1* command in MATLAB. The filter had 50 coefficients and used a Hamming window. In MATLAB, *fir1* command is utilized to digitally design linear-phase FIR filters. The aim of Hamming window technique is to smooth signal edges, diminish abrupt changes, and minimize spectral leakage, ultimately enhancing spectral analysis. This window function gradually transitions from the signal's center to its edges, reducing energy leakage into neighboring frequency bins during Fourier analysis, thereby

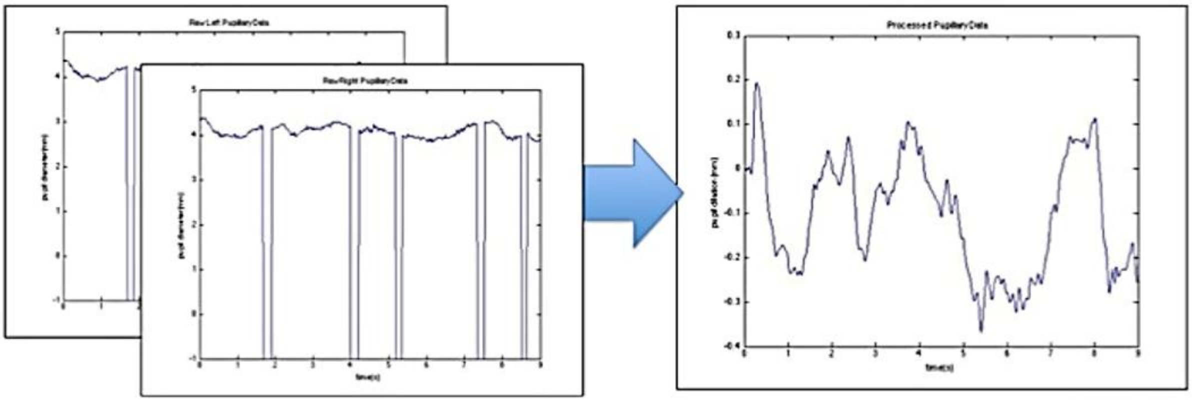

Figure 4. Signal processing: signals with errors (left) & smoothed with Hamming window (right) (Source: Authors)

Table 1. Bin divisions–Three types

Division type #	Description
1	Mean ± standard deviation
2	Mean ± standard error of mean
3	Mean ± midpoint between standard deviation & standard error of mean

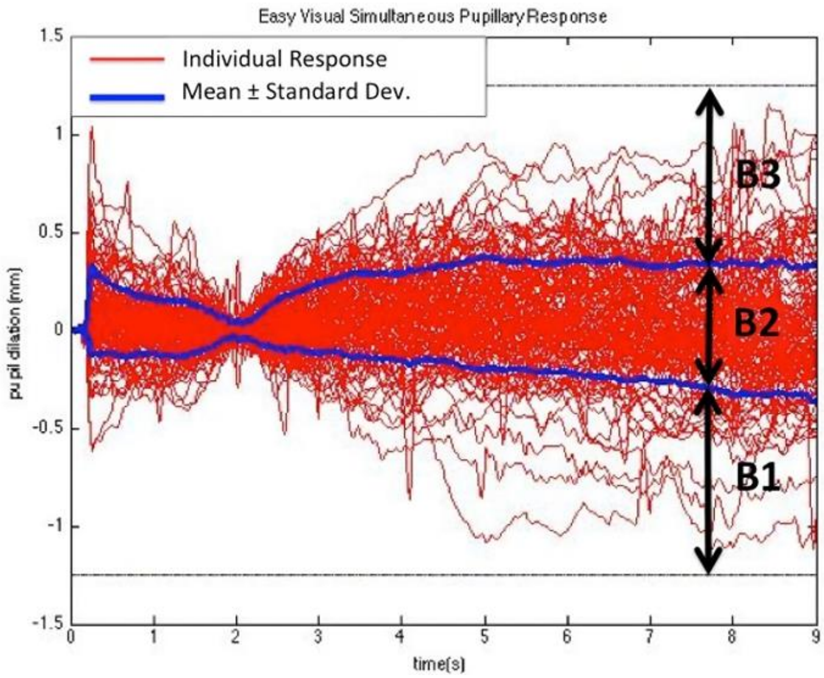

Figure 5. Division type 1 for "easy task" performance pupillary data (B1, B2, & B3, represents bin 1, bin 2, & bin 3, respectively) (Source: Authors)

improving accuracy in frequency representation while maintaining a balance between main and side lobes. The signals from left and right eyes were averages after filtering, which is illustrated in **Figure 4**.

As shown in **Figure 4**, the precision of spectrum analysis is significantly boosted by the Hamming window, which lessens spectral leakage and smoothens signal edges, proving particularly effective when examining finite sections of larger signals with diverse signal characteristics and analytical requirements.

Creation of Bins: Group Analysis

Three types of difficulties and three division types are considered in creating data bins. These three difficulty types are easy, medium, and hard and small, medium, and large are considered as three bins (**Table 1**).

It should be noted that the standard deviation and standard error was taken for the entire signal.

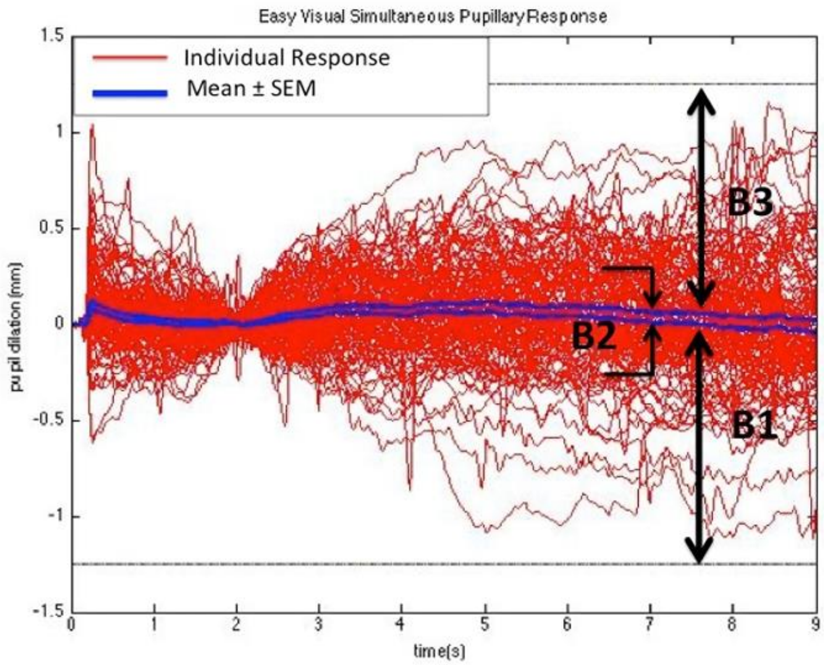

Figure 6. Division type 2 for "easy task" performance pupillary data (B1, B2, & B3, represents bin 1, bin 2, & bin 3, respectively) (Source: Authors)

Figure 5 shows bin type 1 and **Figure 6** shows the bin division type 2. for all trials involving the presentation of an easy multiplication task.

These data were further categorized into three response types including correct, incorrect, and no response for each bin. The resultant bin includes a total of 27 unique response categories. Again, using MATLB each categories' the response frequencies were determined and sorted.

Creation of Bins: Individual Analysis

In cognitive science and human psychology data analysis, response bins refer to distinct categories of data that responses (data points) are grouped for analysis or classification, aiding in simplifying complex data sets. Model parameter descriptions detail the specific attributes, values, or settings within a model that impact its behavior or output, providing crucial information for understanding and fine-tuning the model's performance are shown in detail in the result section. Both response bins and model parameter descriptions are integral components in data analysis and modeling, facilitating individual and group interpretation and optimization of results.

To validate the results of the group analysis, an individual, subject-to-subject analysis was necessary. The individual analysis was conducted on eight out of the 12 subjects. For each subject, the standard deviation of the last 20 samples of the pre-stimulus accommodation period was calculated. This baseline standard deviation was then used to create bin divisions specific to everyone, resulting in the use of five types of bin divisions. At each level of difficulty, the data is sorted into correct, incorrect, and silent (no response) categories for each bin. An illustration of how the bins is divided for an individual pupillary response is presented in **Figure 7**. Five types of bin divisions for an example subject, when presented with an easy MM task, are displayed. The bin divisions are determined by taking mean value plus or minus 3, 5, 7, 10, and 15 standard deviations.

Multinomial Processing Trees

Quantifying latent cognitive states is a significant challenge in developing a cognitive model. To address this, a tree model named MPT is employed as an important analytical tool separating and predicting cognitive processes, with pupil diameter changes. MPT model is a proven technique that describes response probability within a set of possible responses. The tree root represents the presented stimulus, while each branch represents values that are estimated through parameters. The tree may have a different layer representing

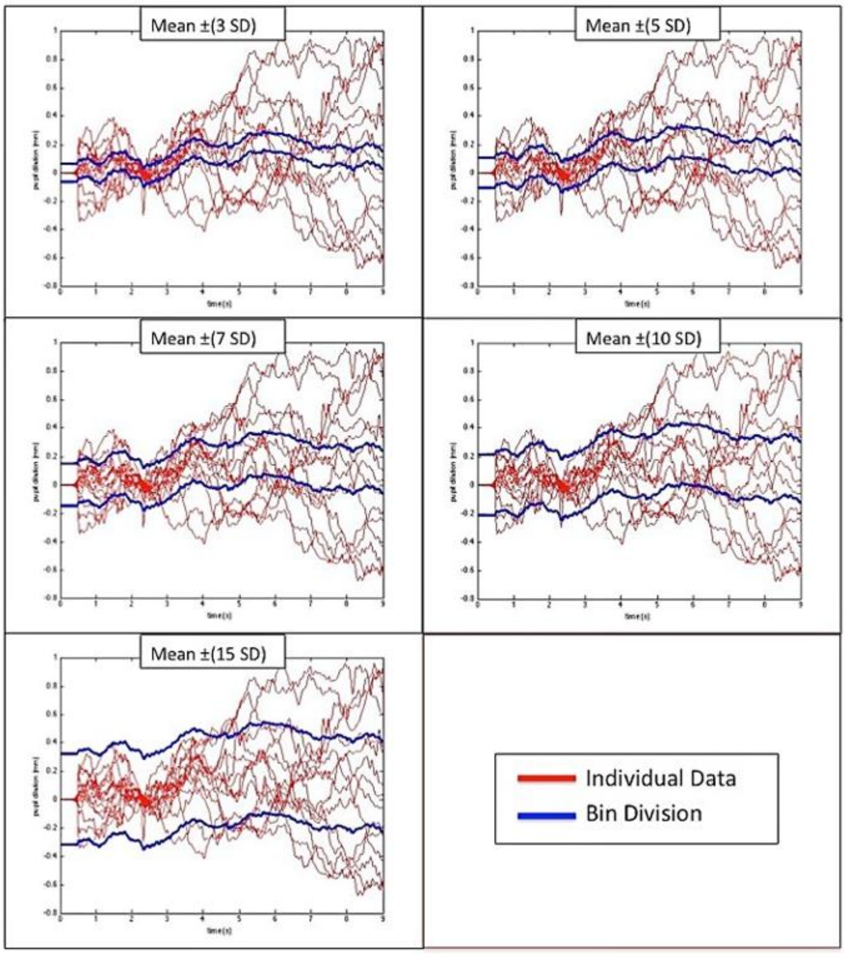

Figure 7. Bin divisions of pupillary response for an example subject when presented an easy mental multiplication task (bin divisions are mean plus or minus 3, 5, 7, 10, & 15 standard deviations) (Source: Authors)

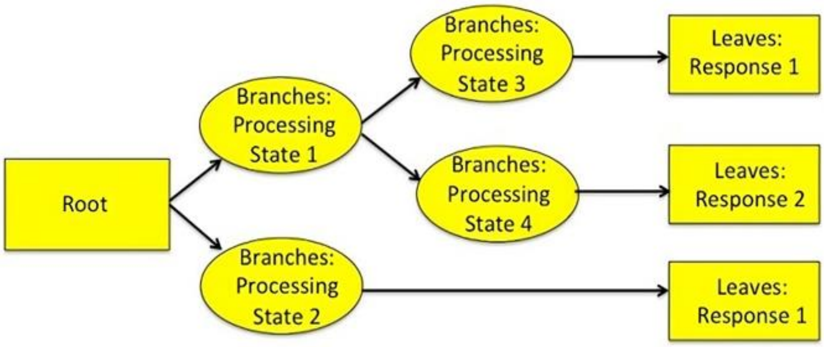

Figure 8. A multinomial processing tree block diagram (Source: Authors)

the probabilities of transitioning from one latent state to the next state. Tree leaves represent the known response frequencies and internal nodes represent hidden layers of the model.

Unlike ANOVA, MPT models decompose the data into essential latent cognitive processes, rather than testing hypotheses on observed data (Singmann, 2010).

Figure 8 shows the structure of an MPT (UMASS, 2014), which is constructed with conditional probabilities known as the fundamental equations and is used in constructing MPTs. **Figure 9** shows an example of such an MPT model, that explains the face recognition process. That is, whether a person can recognize his friend's friend from a recent social gathering. This is a replication of a cued serial recall task experiment, where participants are asked to recall from a list of face pictures (items). The recognition task was defined as, he was

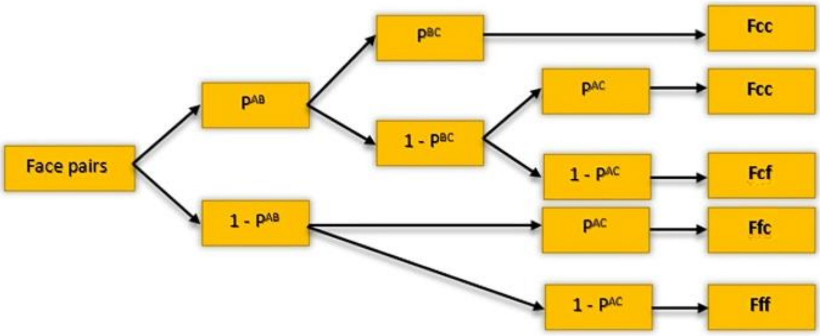

Figure 9. Example of a multinomial processing tree model (Source: Authors)

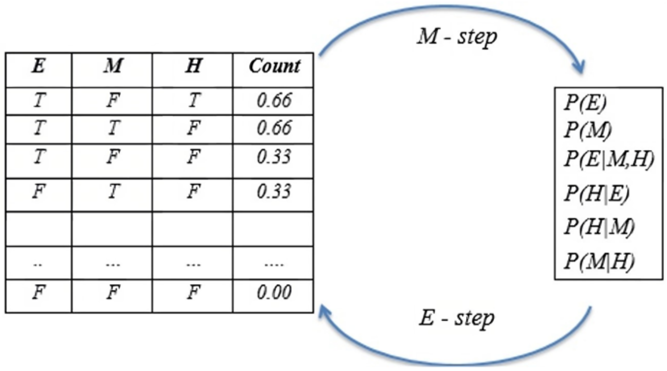

Figure 10. EM algorithm in cognitive task performance modeling (Moshagen, 2010) (E represents 'easy', M represents 'medium', H represents 'hard' task, T represents task performed correctly, & F means task is performed incorrectly) (Source: Authors)

asked the first face (B) and second face (C) that follows the cue face (A). Based on the cognitive ability to recall faces B and/or C, participants' responses are categorized into four groups.

To predict frequencies of occurrences with inter-face image associations between items A, B, and C, a simple model is created. Model considers only forward associations between these faces, requiring three states and, consequently, three parameters: pAB (probability of associations between faces A & B), pBC (probability of associations between faces B & C), and pAC (probability of associations between items A & C).

Depending on the combination of states a person is in, different outcomes would be expected. For example, two sets of states could lead to correct classification of both faces B and C (henceforth referred to as category "Face correctly classified - Fcc"). Additionally, two other possibilities exist:

(1) an association between faces A and B and

(2) an association between faces B and C.

Another possibility is an association between faces A and B, and simultaneously, an association between faces A and C. Both sets of states would result in correct responses for both faces B and C. Therefore, the probability for the category "face correctly classified [FCC]" would be determined based on these combinations of states, as follows:

$$FCC = pAB \times pBC + pAB \times (1 - pBC) \times pAC. \quad (1)$$

In that way the entire model can be described by an equation system, and it can be illustrated by the binary tree with all possible outcomes (face falsely classified [FFC]) model depicted in **Figure 9**.

Out of several software tools for MPT modeling, this paper used multiTree as a software tool for MPT based modeling of pupillary dynamics for cognitive effort estimation (Moshagen, 2010). According to Moshagen (2010), the multitree uses EM algorithm in estimating model parameters. EM model splits the data into two phases, an expectation phase, and a maximization phase, and initializes the parameter at 0.5 value. In "E" phase the previous trial value is used in estimate the expected values. During "M" phase, using values from "E" phase are used in estimating the maximization values. The entire EM model can be represented by

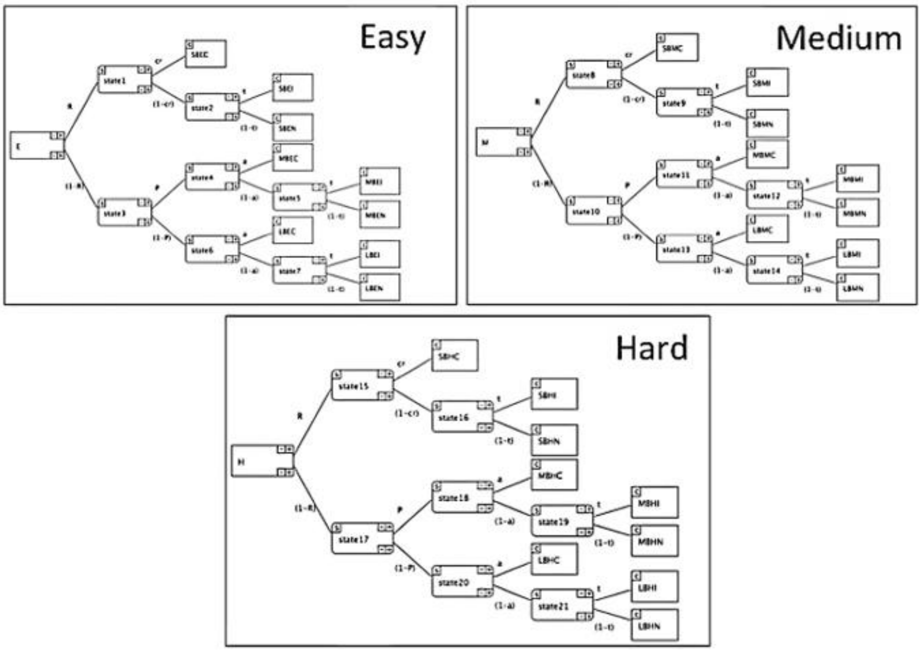

Figure 11. MPT model for individual analysis (easy, medium, & hard) (Source: Authors)

an equation system, which is illustrated using a binary tree depicting all possible outcomes, including "FFC", as shown in **Figure 10**.

A diagram illustrating EM algorithm in task complexity perception is shown in **Figure 10**.

The final model can vary based on model selection criteria that were used to determine the "best" final model. A goodness-of-fit measure G^2 is used in finding discrepancy between the observed data and the statistical predictions. and the observed data without considering the model flexibility. G^2 is defined, as follows:

$$G^2 = 2\sum_{k=1}^{K}\sum_{j=1}^{J_k} n_{j,k}[\ln(n_{j,k}) - \ln(N_k * p_{j,k})], \quad (2)$$

where k is the tree number, K is the total number of trees, j is the branch number, J_k is the total number of branches for a given tree, $n_{j,k}$ is an element of a response frequency vector, N_k is the total number of responses for a given tree, and $p_{j,k}$ is the probability of moving from one branch to a given response category.

Punishment factors are also added to the original G^2 value as part of the Bayesian information criterion (BIC) and Akaike's information criterion (AIC). However, the differences between various models' functional form is ignored in this process, which is one of the drawbacks. AIC and BIC criteria are defined, as follows:

$$AIC = G^2 + 2S, \quad (3)$$

$$BIC = G^2 + \ln(N) \times S, \quad (4)$$

where N is the total number of response and S is the total number of parameters.

To have simpler model the Fisher information approximation (FIA) is used. FIA incorporates all three things, goodness-of-fit, flexibility, and the model's functional form as like the formula:

$$FIA = \frac{1}{2} \times G^2 + \frac{S}{2}\ln\left(\frac{N}{2}\right) + \int \sqrt{\det\left(I(\Theta)\right)}d\Theta, \quad (5)$$

where Θ represents the parameter vector and $I(\Theta)$ is the Fisher information matrix. Some software can be used in tuning the model and parameters (Moshagen, 2010; Singmann, 2010), and minimizing G^2, FIA, AIC, and BIC, which is desirable.

This study can be divided into two phases: The first phase is the best bin division phase–comparing the model selection criteria in the selection of the best bin. Accordingly, based on these response frequencies of each of the three bin division types an MPT model is generated. The second stage is the model-fitting phase, where a model is fitted to the best bin division data.

MPT tree for individual analysis used for this study is shown in **Figure 11**.

Similarly, MPT tree for group analysis used for this study is shown in **Figure 12**.

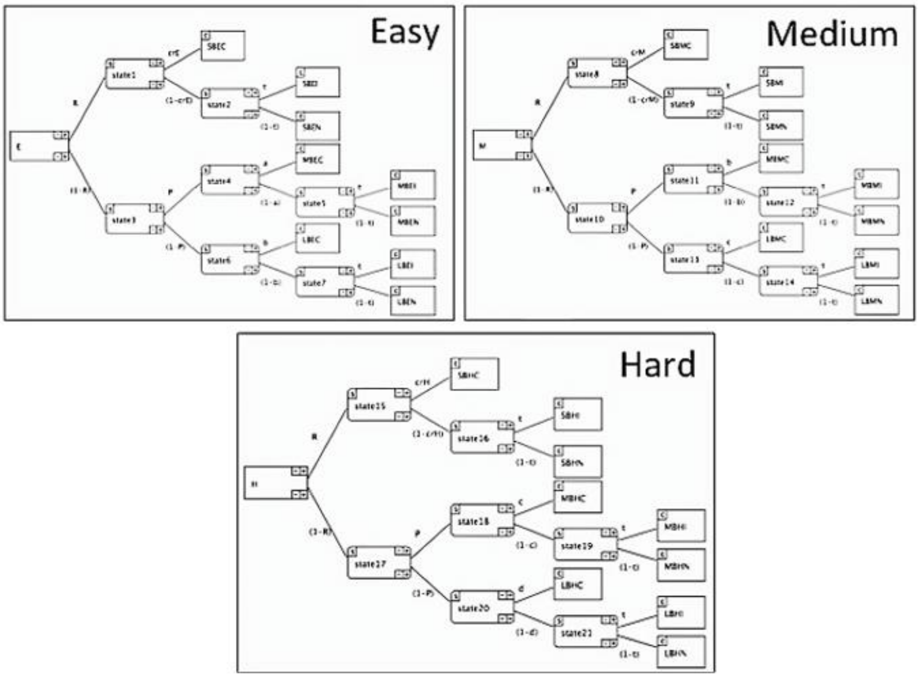

Figure 12. MPT model for group analysis (easy, medium, & hard) (Source: Authors)

Table 2. MPT best bin division phase parameters (both group & individual analyses)

Parameter	Description
R	Probability of recognizing the mental multiplication problem and recalling from LTM
P	Probability of perceiving the problem's level of difficulty
a	Probability of solving the problem
cr	Probability of recalling the correct answer
t	Probability of meeting the time constraints of the task

Table 3. MPT model parameters in **Figure 12** (only for group analysis)

Parameter	Description
R	Best bin division phase
P	Similar to R
a	Perceived probability of solving (easy)
b	Perceived probability of solving (medium)
c	Perceived probability of solving (hard)
d	Solving probability when perceived as extremely difficult
crE	Probability of correct easy recall
crM	Probability of correct medium recall
crH	Probability of correct hard recall
t	Same as R

Table 2 defines MPT model parameter from **Figure 11**.

Table 3 shows MPT model parameter from **Figure 12**.

RESULTS

Analysis results for both the best bin division and the model-fitting phases with respect to both individual and group analyses are explained in this section. **Table 4** and **Table 5** show the comparison results of both phases model selection criteria for each type of bin divisions and the group analysis, respectively.

Group Analysis

Similarly, the model selection criteria for each type of bin division is compared in **Table 4**, and the parameter values of the model-fitting phase are compared in **Table 5**. **Table 4** and **Table 5** include 95% confidence interval (CI) values and standard error margin, and standard error of mean (SEM) values.

Table 4. Model selection criteria–Group comparison

Bin division type #	G^2	AIC	BIC	FIA
1	31.28	1,023	1,063	524.7
2	26.17	1,136	1,175	580.6
3	33.40	1,242	1,281	633.8

Table 5. Model-fitting phase–Group analysis

Parameter	Mean value	SEM	95% lower CL	95% upper CL
R	0.15	0.018	0.11	0.18
P	0.84	0.02	0.80	0.88
a	0.98	0.01	0.95	1.00
b	0.86	0.03	0.80	0.92
c	0.47	0.05	0.34	0.55
d	0.60	0.13	0.352	0.85
crE	0.94	0.06	0.83	1.00
crM	0.92	0.06	0.80	1.00
crH	0.57	0.13	0.31	0.83
t	0.68	0.05	0.59	0.78

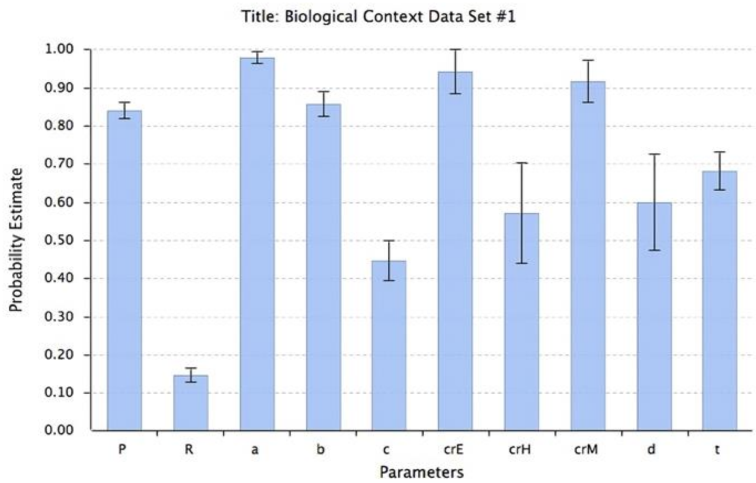

Figure 13. MPT parameter values (with SEM bars) (Source: Authors)

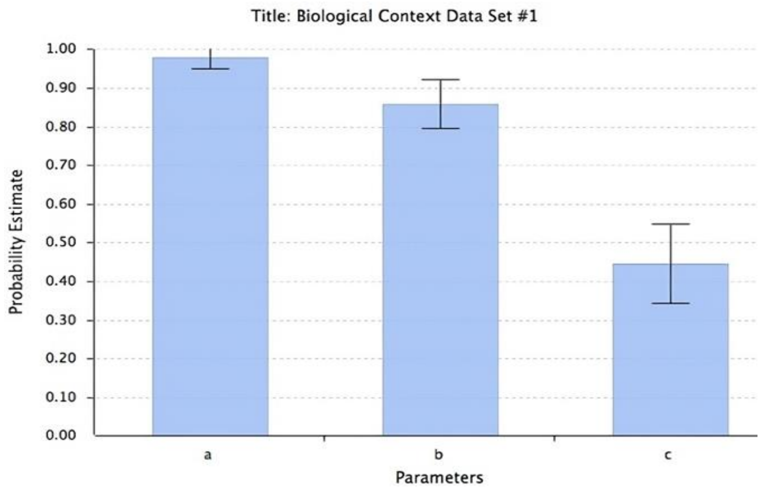

Figure 14. MPT parameter values (with CI) (Source: Authors)

A bar graph displaying all parameter values of interest along with their SEM bars is shown in **Figure 13**.

Additionally, **Figure 14** illustrates a bar graph showing all parameter values with their corresponding CIs.

Notably, the beta distribution is used in sampling and making standard t-tests or ANOVA tests inappropriate due to their normality assumption.

Table 6. Optimal bin divisions for each subject along with corresponding model selection criteria

Subject #	Bin division type	G^2	AIC	BIC	FIA
1	7 SD	22.99	121.60	129.80	64.08
2	3 SD	15.10	107.20	114.10	56.20
3	5 SD	28.62	109.10	116.00	57.17
4	15 SD	9.12	74.32	81.33	39.84
5	5 SD	13.29	103.50	110.50	54.43
6	10 SD	26.71	100.90	107.80	53.06
7	15 SD	10.13	72.09	78.57	38.46
8	3 SD	18.99	93.66	100.49	49.42

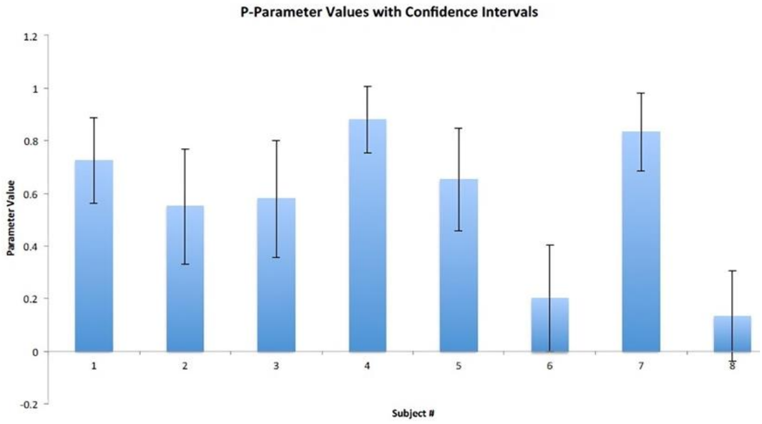

Figure 15. MPT model P parameter values for each subject along with 95% confidence intervals (Source: Authors)

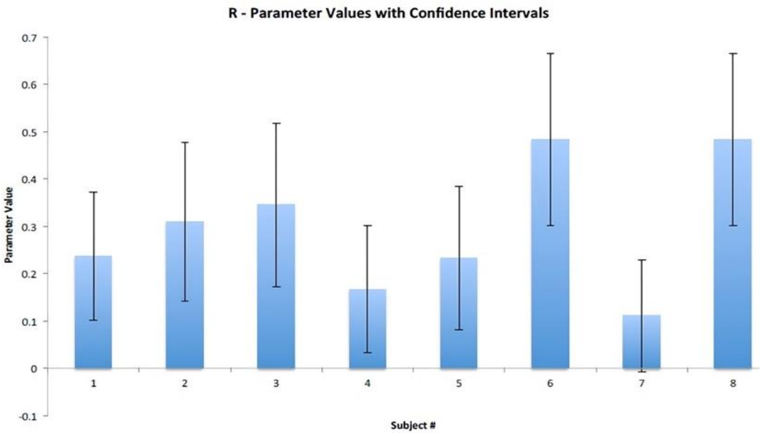

Figure 16. MPT R parameter values for each subject along with 95% confidence intervals (Source: Authors)

The model selection criteria led to the identification of division type 1 as the best bin division type, utilizing the standard deviation of the pupillary response to separate the bins in phase 1.

Significant differences were observed among parameters a, b, and c, as evident from their SEM bars and non-overlapping CIs. This suggests statistical significance, further supported by Kolmogorov-Smirnov test statistics. Based on **Figure 14**, it can be inferred that probability of a subject can solve task decreases with perceived difficulty level. SEM and CI values are used in statistical significance calculation. But future studies should consider implementing a more formal and robust significance test tailored for beta distribution.

Individual Analysis

The model selection criteria for each bin division type for the pupillary response of each individual is compared to determine the optimal bin size for each subject. The optimal bin sizes, along with model selection criteria, are listed in **Table 6**. The parameter values for the optimal bin size model were determined for each subject, and the comparisons of parameters P, R, a, cr, and t across the eight subjects are illustrated in **Figure 15**, **Figure 16**, **Figure 17**, **Figure 18**, and **Figure 19**, respectively.

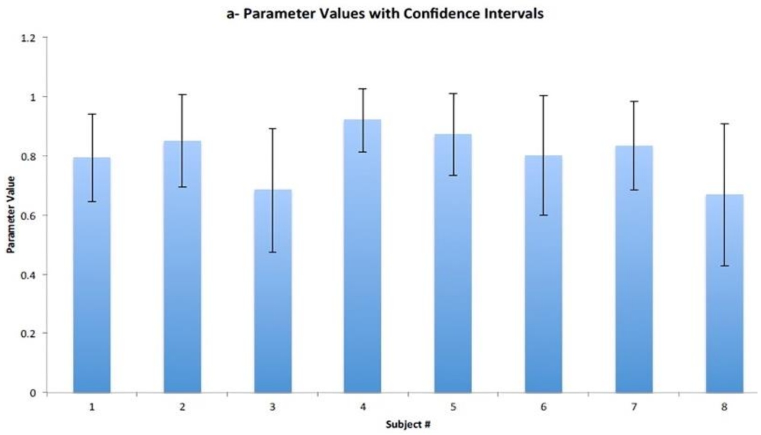

Figure 17. A bar graph of a parameter values for each subject along with 95% confidence intervals (Source: Authors)

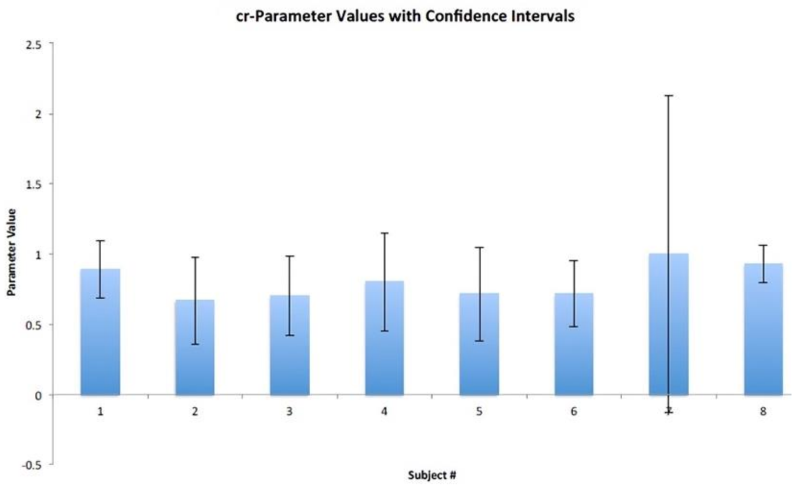

Figure 18. A bar graph of cr parameter values for each subject along with 95% confidence intervals (Source: Authors)

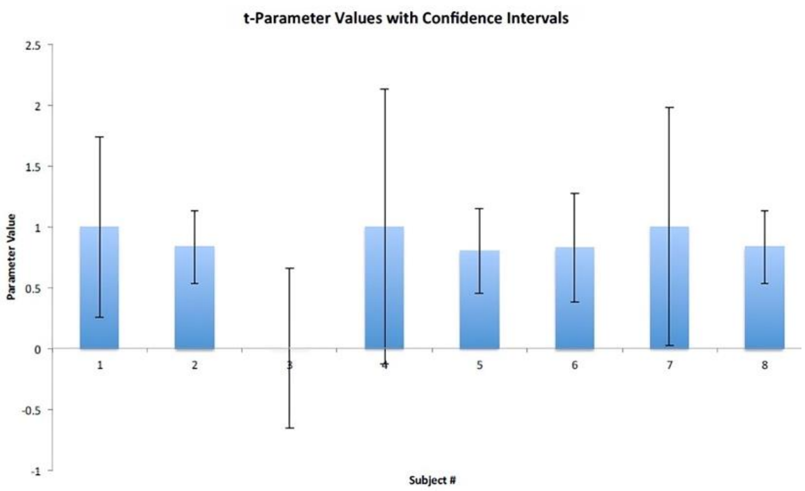

Figure 19. A bar graph of t parameter values for each subject along with 95% confidence intervals (Source: Authors)

The model selection criteria resulted in varying optimal bin sizes from subject-to-subject, primarily due to the high variability observed among the subjects. Following the determination of the optimal bin size for each subject, individual comparisons of each parameter value were made across the eight subjects.

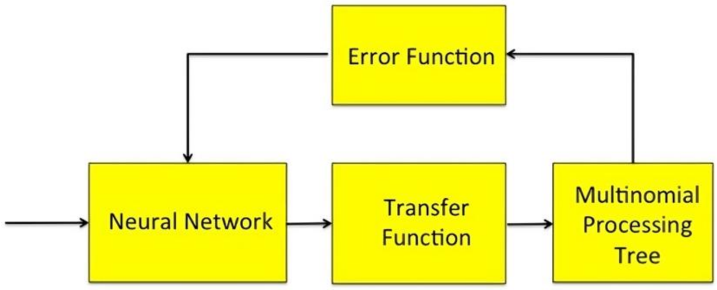

Figure 20. A block diagram of the purpose combined MPT/ANN model (Source: Authors)

Only subject 6 and subject 8 exhibited a significantly different P parameter. Subject 7 differed significantly from subject 7 and subject 8 in the R parameter value. However, no significant differences were found among subjects in the a, cr, and t parameters.

Based on these results, it is suggested that excluding subject 6 and subject 8 from the analysis would lead to a more accurate group analysis. In future work, efforts should be dedicated to performing a group analysis under the same conditions, such as using the same type of standard deviation, for all 12 subjects. This analysis would be based on individual analysis. Additionally, for a more formal evaluation of significance in future studies, a more robust significance test should be implemented.

DISCUSSION

The parameters employed in MPT model primarily possess a cognitive science context. One potential approach to providing a neurobiological context to the model involves connecting MPT with an ANN.

MPT and ANN would constitute two subcomponents of the entire model, interconnected through a transfer function and an error function, as shown in **Figure 20**, based on findings from Chen and Magdon-Ismail (2006).

The parameter values from MPT model would be converted into ANN weights using the transfer function, allowing ANN to learn from MPT parameter values.

The concept of neural gain, defined by Eldar et al. (2013) as the level of norepinephrine released as a function of varying levels of cognitive effort, could be compared between easy, medium, and hard MM tasks using the connected MPT/ANN model. By validating the ANN's connection weights with functional connectivity data from fMRI, the combined model could offer a more cost-effective means of quantifying functional connectivity, considering that fMRI machines are considerably more expensive than eye trackers (Arrington Research, 2015; Block Imaging, 2014).

The pupillary response MPT model, in isolation, has already demonstrated utility in disease diagnostics, as illustrated by O'Neill and Trick's (2001) study comparing pupillary dynamics between narcoleptic and control subjects. The study identified distinct differences in pupillary responses between the two groups, which may affect the resulting pupillary bins. Future research should focus on comparing MPT models between narcoleptic subjects and control subjects.

Using MPT model for diagnosing visual cognitive effort has proven to be a robust approach.

The study's model acts as a concise, cause-and-effect framework, deepening our grasp of how pupils react to mental exertion. Future research aims to craft an MPT that better mirrors the neurological pathways, using parameters to depict the likelihood of nerve impulses passing to subsequent parts of the visual pathway. This advancement could make MPTs a potent tool for diagnosing visual cognitive strain, potentially revolutionizing neurobiological diagnostics.

Advancements in cognitive effort estimation technologies hold the potential to transform our comprehension and engagements with human cognition across diverse contexts. These technologies, as they progress and permeate various sectors, have the capacity to revolutionize driving safety systems, predictive sports performance analytics, human-chatbot interactions, adaptive gaming experiences, as well as advancements in virtual and extended reality for enhanced healthcare innovations such as real-time mental health monitoring. This technological progression stands to significantly benefit numerous modern technologies and applications.

MPT model encounters a limitation in its reliance on the assumption of discrete cognitive processes, posing challenges in accurately estimating cognitive effort. The simplification inherent in the model estimation process may lead to inaccuracies by oversimplifying the intricate nature of cognitive processes. Additionally, MPT models require a priori characterization of cognitive processes, which can be subjective and challenging, impacting the precision of effort estimates. Moreover, when applied in complex real-world scenarios, where cognitive functions are multidimensional and interconnected, MPT algorithm may face limitations in capturing dynamic fluctuations or interactions among cognitive processes.

CONCLUSIONS

MPT model offers a comprehensive integration of various cognitive factors with their respective probabilities of occurrence, allowing for independent treatment of these factors in cognitive task completion. This theoretical basis of computational modeling leads to empirical predictions of cognitive factors related to cognitive load. The benefits of MPT-based cognitive effort estimation are numerous, and some broader predictions can be outlined here. Firstly, the investment of cognitive effort is context-dependent and varies with task complexity. MPT-based predictions can specify context as a core feature in estimating effort and cognitive load, and this can be extended to physical effort estimation, enabling the examination of correlated cognitive and physical exertion. Secondly, by combining MPT model with ANNs, it becomes possible to uncover underlying connectivity between stimulus and action processing areas when subjects engage in high cognitive effort tasks. Moreover, MPT model serves as an important analytical tool in cognitive behavior analysis, bridging neurobiological and psychological analyses. For instance, when dealing with visual cognitive tasks, the model considers the characteristics of human visual pathways (neurobiological) and their effects on behavioral responses (perception-mediated action analysis). By accommodating different neural pathways and performance deficiencies, MPT model offers valuable insights into cognitive load factors. Furthermore, due to the modularity of cognitive effort, MPT modeling opens avenues to investigate cognitively impaired effort processing, considering response time in cognitive processing alongside cognitive source localization in future studies. In summary, MPT model provides a valuable tool for the combined analysis of neurological and psychological effort related to cognitive tasks. It allows for estimating cognitive gain and behavioral cues in cognitively loaded situations, offering a holistic understanding of cognitive effort dynamics.

Author contributions: GH: conceptualized the paper & provided mentorship & **JDE:** performed all data analyses & wrote the first draft of the manuscript. Both authors approved the final version of the article.

Funding: The authors received no financial support for the research and/or authorship of this article.

Acknowledgements: The authors would like to thank Dr. Jeff Klingner of Stanford University for providing raw pupillary data for analysis.

Ethics declaration: The authors declared that the study does not require any ethical approval. The research has no direct interaction with live subjects.

Declaration of interest: The authors declare no competing interest.

Data availability: Data generated or analyzed during this study are available from the authors on request.

REFERENCES

Alsobeh, A., & Shatnawi, A. (2023). Integrating data-driven security, model checking, and self-adaptation for IoT systems using BIP components: A conceptual proposal model. In K. Daimi, & A. Al Sadoon (Eds.), *Proceedings of the 2023 International Conference on Advances in Computing Research* (pp. 533-549). Springer. https://doi.org/10.1007/978-3-031-33743-7_44

Arrington Research. (2015). *Eye tracker prices*. http://www.arringtonresearch.com/prices.html

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin, 91*(2), 276-292. https://doi.org/10.1037/0033-2909.91.2.276

Bishara, A. J., & Payne, B. K. (2008). Multinomial processing tree models of control and automaticity in weapon misidentification. *Journal of Experimental Social Psychology, 45*(3), 524-534. https://doi.org/10.1016/j.jesp.2008.11.002

Block Imaging. (2014). Your guide to medical imaging equipment. https://info.blockimaging.com/success-stories

Böhm, M. F., Bayen, U. J., & Schaper, M. L. (2020). Are subjective sleepiness and sleep quality related to prospective memory? *Cognitive Research: Principles and Implications*, *5*(1), 5. https://doi.org/10.1186/s41235-019-0199-7

Chen, H.-C., & Magdon-Ismail, M. (2006). NN-OPT: Neural network for option pricing using multinomial processing tree. In I. King, J. Wang, L. W. Chan, & D. Wang (Eds.), Neural information processing (pp. 360-369). Springer. https://doi.org/10.1007/11893295_41

Cross Check Networks. (2015). *SOA testing techniques*. http://www.crosschecknet.com/soa_testing_black_white_gray_box.php

da Silva Castanheira, K., LoParco, S., & Otto, A. R. (2021). Task-evoked pupillary responses track effort exertion: Evidence from task-switching. *Cognitive, Affective, & Behavioral Neuroscience, 21*, 592-606. https://doi.org/10.3758/s13415-020-00843-z

de Gee, J. W., Knapen, T., & Donner, T. H. (2014). Decision-related pupil dilation reflects upcoming choice and individual bias. *PNAS, 111*(5), E618-E625. https://doi.org/10.1073/pnas.1317557111

Einhäuser, W., Stout, J., Koch, C., & Carter, O. (2008). Pupil dilations reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *PNAS, 105*(5), 1704-1709. https://doi.org/10.1073/pnas.0707727105

Eldar, E., Cohen, J. D., & Niv, Y. (2013). The effects of neural gain on attention and learning. *Nature Neuroscience, 16*, 1146-1153. https://doi.org/10.1038/nn.3428

Griffiths, T. L., & Kalish, M. L. (2001). A multidimensional scaling approach to mental multiplication. *Memory and Cognition, 30*(1), 97-106. https://doi.org/10.3758/BF03195269

Haines, D. E. (2013). *Fundamental neuroscience for basic and clinical applications*. Elsevier.

Hossain, G., & Elkins, J. D. (2018). When does an easy task become hard? A systematic review of human task-evoked pupillary dynamics versus cognitive efforts. *Neural Computing and Applications*, *30*, 29-43. https://doi.org/10.1007/s00521-016-2750-5

Jarrah, A., Almomany, A., Alsobeh, A. M. R., & Alqudah, E. (2021). High-performance implementation of wideband coherent signal-subspace (CSS)-based DOA algorithm on FPGA. *Journal of Circuits, Systems and Computers, 30*(5), 2150196. https://doi.org/10.1142/S0218126621501966

Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes exploration-exploitation trade-off: Evidence for the adaptive gain theory. *Journal of Cognitive Neuroscience, 23*(7), 1587-1596. https://doi.org/10.1162/jocn.2010.21548

Klingner, J. (2010). *Measuring cognitive load during visual tasks by combining pupillometry and eye tracking* [Doctoral dissertation, Stanford University].

Klingner, J., Tversky B., & Hanrahan, P. (2011). Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks. *Psychophysiology, 48*(3), 323-332. https://doi.org/10.1111/j.1469-8986.2010.01069.x

Köstering, L., McKinlay, A., Stahl, C., & Kaller, C. P. (2012). Differential patterns of planning impairments in Parkinson's disease and sub-clinical signs of dementia? A latent-class model-based approach. *PLoS ONE, 7*(6), e38855. https://doi.org/10.1371/journal.pone.0038855

Miles, A., Brett, G., Khan, S., & Samim, Y. (2023). Testing models of cognition and action using response conflict and multinomial processing tree models. *Sociological Science, 10*, 118-149. https://doi.org/10.15195/v10.a4

Moshagen, M. (2010). multiTree: A computer program for the analysis of multinomial processing tree models. *Behavior Research Methods, 42*, 42-54. https://doi.org/10.3758/BRM.42.1.42

Nestler, S., & Erdfelder, E. (2023). Random effects multinomial processing tree models: A maximum likelihood approach. *Psychometrika, 88*, 809-829. https://doi.org/10.1007/s11336-023-09921-w

Nowak, W., Hachol, A., & Kasprzak, H. (2008). Time-frequency analysis of spontaneous fluctuations of the pupil size of the human eye. *Optica Applicata* [*Applied Optics*], *38*(2), 469-480.

O'Neill, W. D., & Trick, K. P. (2001). The narcoleptic cognitive pupillary response. *IEEE Transactions on Biomedical Engineering, 48*(9), 963-968. https://doi.org/10.1109/10.942585

Ohtsuka, K., Asakura, K., Kawasaki, H., & Sawa, M. (1988). Respiratory fluctuation of the human pupil. *Experimental Brain Research, 71*, 215-217. https://doi.org/10.1007/BF00247537

Privitera, C. M., Renninger, L. W., Carney, T., Klein, S., & Aguilar, M. (2010). Pupil dilation during visual target detection. *Journal of Vision, 10*(10), 3. https://doi.org/10.1167/10.10.3

Reilly, J., Kelly, A., Kim, S. H., Jett, S., & Zuckerman, B. (2019). The human task-evoked pupillary response function is linear: Implications for baseline response scaling in pupillometry. *Behavior Research Methods, 51*, 865-878. https://doi.org/10.3758/s13428-018-1134-4

Sensi, F., Calcagnini, G., & De Pasquale, F. (1999). Baroreceptor-sensitive fluctuations of human pupil diameter. *Computers in Cardiology, 1*, 233-236.

Singmann, H. (2010). MPTinR: Analysis of multinomial processing trees in R. *Behavior Research Methods, 45*, 560-575. https://doi.org/10.3758/s13428-012-0259-0

Stergiou, C., & Siganos, D. (1996). *Neural networks*. https://srii.sou.edu.ge/neural-networks.pdf

UMASS. (2014). *Multinomial processing tree models*. http://people.umass.edu/alc/course_pages/fall_2004/modeling_behavior/lectures/MPTs.ppt