

# Single Trial Variability of Event-Related Brain Potentials as an Index of Neural Efficiency During Working Memory

David W. Shucard<sup>(✉)</sup>, Thomas J. Covey, and Janet L. Shucard

Division of Cognitive and Behavioral Neurosciences, Department of Neurology, Neuroscience Program, The Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, The State University of New York, Buffalo, USA  
{dshucard, tjcovey, shucard}@buffalo.edu

**Abstract.** Event-related brain potentials (ERPs) to a particular stimulus are extracted from the continuous electroencephalogram (EEG) through signal averaging techniques. The most extensively studied ERP component, P300 (P3, or P3b), occurs at approximately 300–800 ms post-stimulus. P3 amplitude and latency are markers of the attentional/cognitive resources devoted to the task and the timing (within msec) of central processing speed of an individual's cognitive response to a stimulus, respectively. Variability in the timing and amplitude of components in the single EEG trials that contribute to the averaged ERP has been of particular interest to our laboratory because it may provide an index of central information processing efficiency. Examination of single trial variability (STV) can provide a level of analysis beyond traditional ERP measures and offers a unique marker of the functional integrity of neural pathways. In the present study we examined ERP STV as it relates to WM demand or load during a visual n-back task in normal adult participants. Performance measures and the scalp-recorded EEG were obtained during the n-back task. Frontal and parietal scalp averaged ERPs were derived from the EEG time-locked to the stimuli. The mean and variability (SD) of the peak amplitude and latency were then obtained from the single trial data for each participant and condition. Results showed that as WM load increased, behavioral measures of processing speed slowed, behavioral efficiency decreased, and the number of correct responses decreased. Correlations for both latency and amplitude between the P3 component derived from the averaged ERP and the P3 component derived from the single trials were generally high, indicating that the P3s for individual trials identified by our STV procedure were representative of the P3s in the averaged ERP obtained by standard signal averaging procedures. P3 STV analyses also showed differential effects between frontal and parietal scalp sites for both amplitude and latency variability that were related to WM load. Both frontal P3 latency STV and amplitude increased as WM load increased, indicating decreased neural efficiency associated with an increase in WM load. Single trial ERP variability measures may provide potential physiological markers of the neural efficiency of brain processes engaged in cognitive functions, such as working memory.

**Keywords:** Working memory · ERP · Single trial variability

# 1 Introduction

## 1.1 Electrophysiological (Event-Related Potentials, ERPS)

Event-related brain potentials (ERPs) to a particular stimulus (e.g., letters that are designated as “targets”) are extracted from the continuous electroencephalogram (EEG) through signal averaging techniques. The most extensively studied ERP component, P300 (P3, or P3b), occurs approximately 300 to 800 ms post-stimulus, and its latency and amplitude are influenced by cognitive factors such as the ability to detect a target and the salience of the target to the subject [see for review 1, 2]. The P3 amplitude and latency are markers of the attentional resources devoted to the task and the timing (central processing speed) of a subject’s cognitive response to a stimulus, respectively. To this end, the P3 component is ideal for studying the physiological correlates of attention in both normal and clinical populations, when difficulties in attention, speed of information processing, and working memory may be present.

## 1.2 The ERP and Working Memory

Working memory (WM), or the ability to maintain and manipulate information briefly in time, is key to higher cognitive functioning such as language, planning, and problem solving [see for example, 3]. Working memory is conceptualized as a network of distributed systems invoking higher order attentional demands required for holding and manipulating information in one’s mind for a relatively short period of time. Working memory has been divided into two processes: Executive control (responsible for the encoding, manipulation, and retrieval of information) and maintenance (keeping information available during a period of time). These working memory processes, mainly through fMRI studies, have been associated with specific brain structures: The prefrontal cortex (responsible for executive control processes and some role in active maintenance), and the posterior areas including the parietal cortex (responsible for active maintenance) [4–7] are two of the primary brain areas involved in WM. Druzgal and D’Esposito [8] found that during encoding and delay (maintenance) periods, prefrontal activation increased with memory load. We found significant WM deficits in patients with Systemic Lupus Erythematosus (SLE) that increased with WM load [e.g., 9].

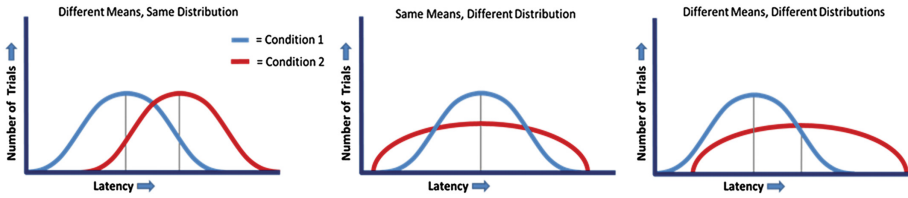
Although imaging techniques lend themselves well to the identification of the anatomical components of WM, for example, the temporal sequence of events is best studied using ERPs. ERPs provide high temporal resolution (in milliseconds), indices of resources allocated to the task, and scalp topographical representations. It is possible to obtain a precise measure of the timing of activity induced by frontal cortical networks activated during cognitive processes such as the different components of WM. A study by McEvoy et al. [10] measured ERPs to a verbal and spatial n-back task and found that increasing the WM demands (or load) resulted in attenuation of P3 amplitude at the parietal site. Additionally, the correct “matches” produced greater amplitude P3 than nonmatching stimuli. They suggested that their findings can be interpreted in terms of the amount of neural resources allocated to the task. A number of other studies reported similar findings of a posterior P3 amplitude decrease as WM load increased, as well as

a frontal P3 amplitude increase as a function of increased load [11, 12]. Research in our laboratory suggested that there is a balance between frontal and parietal brain regions related to the allocation of attentional resources and WM load, measurable by ERP P3 amplitude [13].

### 1.3 Single Trial Variability

Recently, there has been increased interest in intra-individual variability in both the behavioral and neuroscience literature [14, 15]. Intra-individual variability is considered to play a significant role in information processing capacity in the individual. For example, Bielak et al. reported that intra-individual variability of reaction time (RT) predicted cognitive outcomes five years later. Measures of intra-individual variability are of interest because they provide an index of central processing efficiency. Previously, we have used measures of single trial ERP latency variability to study mechanisms underlying ERP average amplitude changes in animals [see for example, 16, 17]. The amount of variability in ERP peak latency across single trials has an effect on the averaged ERP response such that increased STV can decrease the averaged peak amplitude and increase its latency. Such information about the single trials that comprise the ERP average can provide another level of analysis of ERP data besides traditionally used measures of ERP component amplitude and latency described above. In the present study, analysis of STV of the P3 response offered further insight into possible mechanisms responsible for the P3 amplitude and latency differences seen during different cognitive conditions and more importantly, STV could provide a functional marker of the integrity of neural pathways related to cognitive processes in different clinical populations that have cognitive dysfunction such as Multiple Sclerosis. For example, white matter integrity might be compromised in certain patients, leading to greater STV and, in turn, deficits in processing speed and WM.

Thus, variability in the timing and amplitude of ERP components in the single EEG trials that contribute to the averaged ERP could serve as an index of central processing efficiency and has the potential to provide unique information about the recruitment of brain resources for complex cognitive processes such as WM. Note that the distributions of a single trial ERP component illustrated in Fig. 1 could also potentially be independent of the average of the ERP component. In the figure to the left, there is a longer mean latency for Condition 2 and no difference in STV between Condition 1 and 2. In the middle figure, the mean latencies are identical between Conditions even though there is a broader distribution of single trials for Condition 2. In the figure on the right, there is longer mean latency and greater STV for Condition 2 than Condition 1. Thus, the average ERP latency/amplitude could be independent of or dependent on the STV. The relationship between the average and the single trials, we believe is important to determine because it could provide information about the efficiency of the system. For example, the relationship between single trials and the average may look like the figure on the right in patients with white matter damage or demyelinating processes.



**Fig. 1.** Hypothetical frequency distribution of single trials for the latency of the P3 ERP component. (Color figure online)

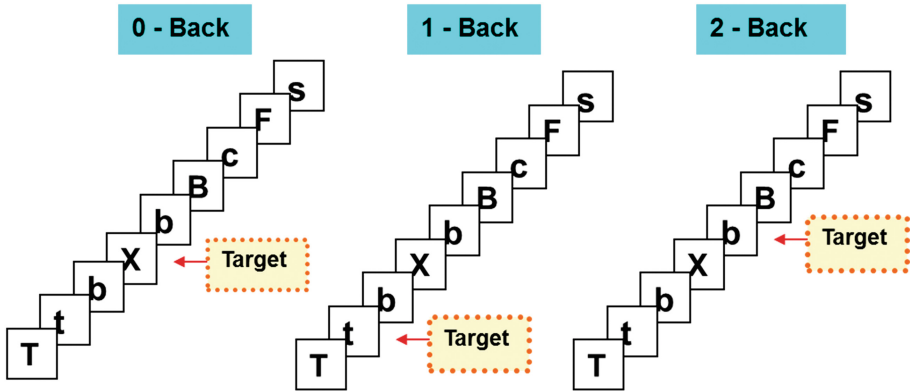
Using a program developed in our laboratory, we examined ERP P3 STV (both amplitude and latency) as it relates to WM load during a visual n-back task in normal adult participants. We predicted that STV of P3 amplitude and latency would increase as a function of the load placed on WM, reflecting increased variability in central processing due to increased demands on WM.

## 2 Methods

**Participant Characteristics:** Twenty healthy controls were tested [age in years, mean = 45.7, SD = 9.33; 65 % female; years of education, mean = 14.2, SD = 1.77; estimated IQ (National Adult Reading Test, NART), mean = 108, SD = 8.07]. This study was approved by an internal Institutional Review Board of the State University of New York at Buffalo and written informed consent was obtained from all participants.

**Visual n-back Paradigm:** The visual n-back working memory task, illustrated in Fig. 2, was used to test processing speed and WM. The n-back is a data rich task that allowed us to obtain behavioral and electrophysiological data simultaneously while manipulating WM load. The behavioral data derived from the n-back task included information processing speed (as measured by RT to both match and non-match stimuli) and accuracy of the response. ERP data were obtained under the three WM load conditions of increasing difficulty (n-back conditions, 0-, 1-, 2-back, consecutively). The n-back task requires attention to visual stimuli and the ability to remember and determine the presence of a particular target stimulus. For the 0-back condition, participants were asked to identify a target letter “X” in a sequence of letters presented one at a time on a computer screen. For the 1-back condition, participants determined whether or not each letter matched the previously presented letter (one back). For the 2-back condition, participants determined whether or not each letter matched the letter two letters back. There were 150 trials in each condition. Stimulus duration was 400 ms with a 2 s inter-stimulus interval. Eleven letters of the alphabet were used as stimuli. Participants were required to respond via button presses to all trials (match and non-match).

**ERP Analysis:** Continuous EEG was recorded during the n-back procedure using a Neuroscan EEG data acquisition and analysis system. Data were recorded from 14 scalp sites with a linked ears reference and forehead ground. Data were filtered during acquisition with a bandpass of 0.1–100 Hz and digitized at a sampling rate of 250 Hz.



**Fig. 2.** Diagram of the visual n-back task

EEG data were segmented epochs with 300 ms pre-stimulus and 1600 ms post-stimulus. A 0.1–25 Hz filter was applied with 24 dB octave roll off, and automated artifact rejection procedures were conducted. Averaged ERPs were obtained for each participant for each condition. Because of the theoretical significance of the frontal-parietal network in WM function, we focused our averaged and STV P3 analyses on frontal and parietal midline scalp leads: Fz and Pz, for 0-, 1-, and 2-back conditions.

**Single Trial Variability Analysis:** Grand averaged ERPs for each condition/lead served as templates for the determination of a window surrounding the P3 component. Figure 3 presents a single trial analysis for an individual participant and the averaged ERP resulting from single EEG trials. The window width was defined by taking the average distance (in msec) from the N2 (negative ERP component occurring before the P3) to the P3 component for each of the grand averaged ERPs. An algorithm written in MATLAB selected the P3 for each trial by determining the region in the window of the single trial that most strongly correlated (point by point) with the template ERP P3 component. Single trial amplitudes that were below  $-25 \mu\text{V}$  or above  $70 \mu\text{V}$  were excluded from the STV analyses. Non-baseline corrected data were used. Figure 3 also illustrates the distribution of P3 latencies and amplitudes across trials for a participant as determined by the STV analysis. Note the variability of peak latencies and amplitudes across trials.

The averaged ERP at Fz for each n-back condition is presented across the top of the figure. Just below the averages is an overlay of the individual trials that make up each averaged ERP (single trial EEG). The bottom two rows of the figure are graphic illustrations of the variability of the P3 response detected by the program. Note that STV of latency increases as WM load increases from the 0-back to the 2-back.

**Statistical Analyses:** Pearson correlations were used to assess the relationships between various measures. Repeated measures ANOVA were used to evaluate condition and lead effects. Post hoc testing was conducted using paired t-tests. Significance was set at  $p \leq .05$ .

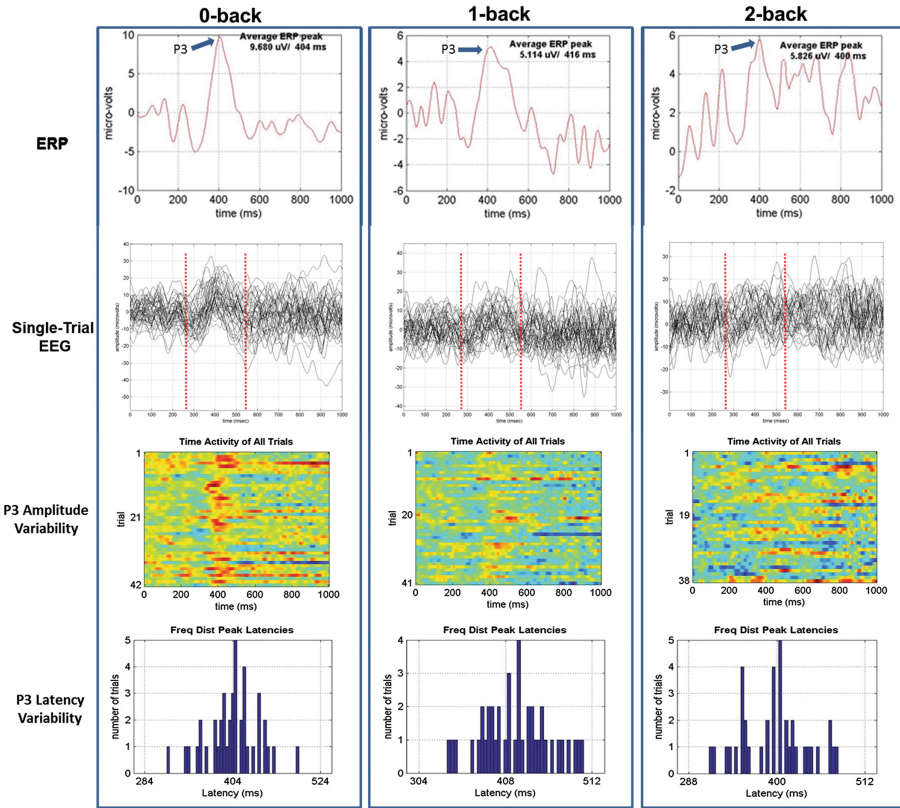


Fig. 3. Averaged ERPs and single trial data from one electrode for one participant

### 3 Results

**Reliability of the Measure:** Table 1 below presents the relationships obtained between the ERP template and the single trial P3 components identified for the frontal and parietal scalp leads across all participants, within the pre-defined component window. Note that the coefficients were all good accounting for about 50 % of the variance [e.g.,  $(0.76^2) = .58$ ]. These findings indicate that the template selected from the participant’s averaged ERP was successfully used to identify the P3 component in the single trials for most participants.

In addition to the relationship between the template window derived for the P3 component and the P3 identified in the single trials, we examined the relationship between the participant’s P3 amplitude and latency derived from the single trial analysis, averaged across the individually selected trials, and the P3 amplitude and latency obtained via standard signal averaging (component selected following averaging). Tables 2 and 3 present these data. Note that the correlation coefficients are generally high for both amplitude and latency, however there appears to be weaker relationships between the derived and standard latencies than with amplitude, particularly for the parietal lead during the condition with the highest WM load (2-back).

**Table 1.** Correlation coefficients between the P3 template and the single trial P3 response

	Mean	St. Dev.	Min	Max
Fz 0-back	0.76	0.04	0.67	0.82
Pz 0-back	0.74	0.06	0.55	0.81
Fz 1-back	0.73	0.04	0.66	0.79
Pz 1-back	0.70	0.05	0.65	0.80
Fz 2-back	0.72	0.05	0.60	0.79
Pz 2-back	0.71	0.04	0.65	0.77

**Table 2.** Correlations between ERP P3 amplitudes derived from single trial analysis and those from the averaged ERP

Derived avg P3 amplitude	Event-related potential P3 amplitude					
	Fz 0-back	Pz 0-back	Fz 1-back	Pz 1-back	Fz 2-back	Pz 2-back
Fz 0-back	.942***	-	-	-	-	-
Pz 0-back	-	.942***	-	-	-	-
Fz 1-back	-	-	.940***	-	-	-
Pz 1-back	-	-	-	.952***	-	-
Fz 2-back	-	-	-	-	.905***	-
Pz 2-back	-	-	-	-	-	.949***

\*\*\* =  $p < .001$   
 \*\* =  $p < .01$

**Table 3.** Correlations between ERP P3 latencies derived from the single trial analysis and those obtained from the averaged ERP

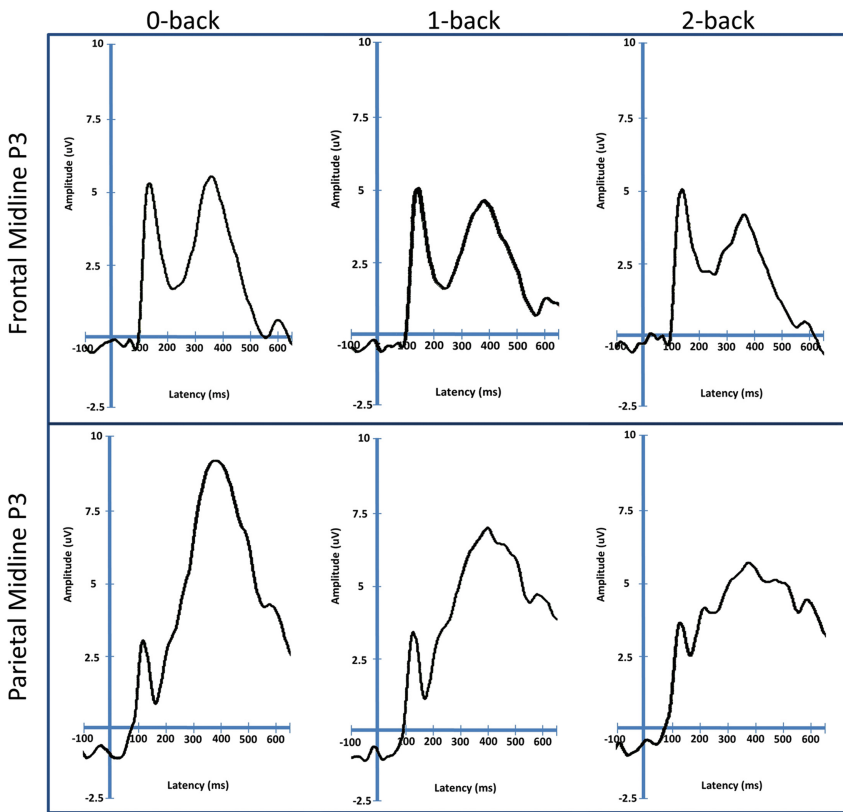
Derived avg P3 latency	Event-related potential P3 latency					
	Fz 0-back	Pz 0-back	Fz 1-back	Pz 1-back	Fz 2-back	Pz 2-back
Fz 0-back	.957***	-	-	-	-	-
Pz 0-back	-	.776***	-	-	-	-
Fz 1-back	-	-	.955***	-	-	-
Pz 1-back	-	-	-	.956***	-	-
Fz 2-back	-	-	-	-	.872***	-
Pz 2-back	-	-	-	-	-	.567***

\*\*\* =  $p < .001$   
 \*\* =  $p < .01$

**Behavioral Effects:** In that the focus of this study was on the STV measure, the performance effects will only be summarized here. Accuracy (correct matches), information processing speed (Reaction Time, RT), and behavioral efficiency (Standard

deviation of RT), all were related to WM load. That is, as WM load increased, performance decreased (ANOVA yielded significant Condition effects for each of these measures). Accuracy decreased particularly between the 1 and 2-back conditions. Reaction time and standard deviation of RT increased from the 0 to 1 to 2 back conditions.

**ERP Findings:** A Condition  $\times$  Lead ANOVA for P3 latency failed to yield any significant effects, although P3 latency at the Pz scalp site was generally longer across all three n-back conditions. A Condition  $\times$  Lead ANOVA for P3 amplitude yielded a significant interaction (Condition  $\times$  Lead,  $F = 4.25$ ,  $p = .025$ ). The findings indicate that P3 amplitude was highest for Pz at the 0-back condition and decreased as a function of WM load. The P3 amplitude for Fz was lower than Pz but did not show a significant decrement from the 1- to 2-back condition. Figures 4 and 5 illustrate these results.



**Fig. 4.** Grand averaged ERPs to the 0-, 1-, and 2-back conditions



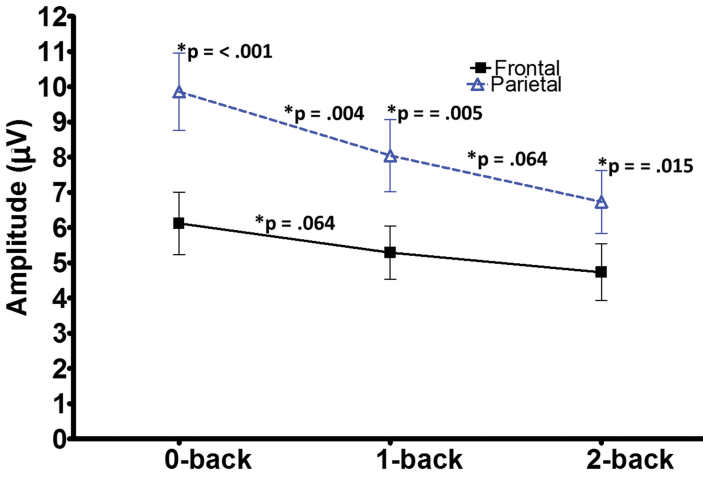


Fig. 5. Condition × Lead effect for P3 ERP amplitude

**Single Trial Variability Findings:** Separate Condition × Lead ANOVAs with STV as the dependent measure for ERP latency and amplitude yielded significant Condition and Lead effects for latency ( $F = 7.40, P = .005$ ;  $F = 8.36, P = .009$ , respectively) and a significant Condition × Lead interaction for Amplitude ( $F = 4.80, p = .015$ ). Figure 6a–c illustrate these findings. As can be seen from Fig. 6B and C, P3 latency variability across trials was greatest during the 1- and 2-back conditions compared to the 0-back. The 2-back condition had greater latency STV than either the 0- or 1-back. In regards to lead, STV was greatest at the parietal scalp site compared to the frontal site. As seen in Fig. 6A, the pattern of amplitude variability across n-back conditions differed between Fz and Pz scalp sites. Importantly, frontal P3 STV for amplitude markedly increased from the 1- to the 2-back condition. The parietal P3 did not show this effect.

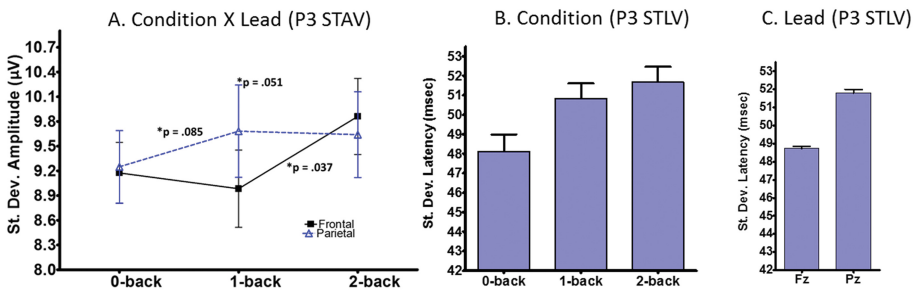


Fig. 6. STAV = single trial amplitude variability; STLV = single trial latency variability.

## 4 Discussion

In this study we examined the efficacy of using the variability of the single trials that comprise an average ERP as a biomarker of neural efficiency during cognition. We examined P3 STV as it relates to WM load during a visual n-back task in normal adult participants. We predicted that STV of P3 amplitude and latency would increase as a function of the load placed on WM, reflecting increased variability in central processing due to increased demands on WM. We showed that n-back behavioral performance varied as a function of WM load. Performance accuracy decreased, information processing speed as measured by RT became slower, and variability of RT, a behavioral index of neural efficiency increased with increased WM load. As reported in the literature, P3 amplitude decreased, particularly at the Pz scalp site as a function of WM load. Interestingly, STV for P3 latency increased with WM demand, despite the fact that there were no condition effects for the averaged ERP P3 latency. This finding indicates that increased STV of P3 latency across trials may occur independently of whether the signal averaged ERP P3 latency measure changes. With minimal demand on WM (0-back) STV for P3 amplitude at Fz and Pz was similar. At moderate WM demand (1-back), there was a subtle (albeit, not significant) increase in STV of P3 at Pz. With high WM demand (2-back), STV of P3 at Fz significantly increased from that seen in the 1-back condition. These STV amplitude findings could indicate differential recruitment of frontal and parietal resources during different levels of WM engagement. During moderate WM demand, increased variability at Pz could reflect engagement of posterior maintenance processes. During the 2-back, there is likely increased recruitment of neuronal resources associated with the frontal cortex due to the need for more executive or top down control, necessitated by increased task demand. Increased STV amplitude of P3 at Fz may reflect recruitment of these frontal resources. Placed within the context of the literature that relates frontal and parietal regions to executive and maintenance aspects of WM, respectively, the results are consistent with the possibility that Pz P3 amplitude variability may reflect maintenance, whereas Fz P3 amplitude variability may reflect central executive demands during WM.

Thus, P3 STV provided information that was not apparent in the averaged ERP P3 component. Aside from the present one, studies have shown that there is variability in ERP peak latency across single trials, and this variability affects the averaged ERP response. In general, higher single-trial-latency variability leads to lower averaged ERP amplitude [16]. Thus, information about the single-trials that comprise the ERP average can provide another level of analysis of ERP data [18]. For example, because trial-to-trial latency variability represents change over time, it can be used to explore cortical dynamics [19] which otherwise would be lost with conventional averaging. To further test the efficacy of single trial ERP variability as a biomarker of neural efficiency, studies should be conducted in which ERP measures are linked with MRI measures of microscopic damage such as demyelinating processes in neural networks involved in cognition. Such structure-function studies will help clarify the functional significance of the variability of single trials that comprise the averaged ERP.

**Acknowledgements.** Partial support for this research was provided by a grant from the National Multiple Sclerosis Society and the Jog for the Jake Foundation.

## References

1. Polich, J.: Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* **118**, 2128–2148 (2007)
2. Kok, A.: On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology* **38**, 557–577 (2001)
3. Baddeley, A.: Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* **4**, 829–839 (2003)
4. Carlson, S., Martinkauppi, S., Rama, S., Salli, P., Korvenoja, E., Aronen, H.J.: Distribution of cortical activation during visuospatial n-back tasks as revealed by functional magnetic resonance imaging. *Cereb. Cortex* **8**(8), 743–752 (1998)
5. Cohen, J.D., Braver, W.M., Nystrom, T.S., Noll, L.E., Jonides, D.C., Smith, J.: Temporal dynamics of brain activation during a working memory task. *Nature* **386**, 604–607 (1997)
6. Fuster, J.M.: Frontal lobes. *Curr. Opin. Neurobiol.* **3**, 160–165 (1993)
7. Gathercole, S.E.: Neuropsychology and working memory. *Neuropsychology* **8**, 494–505 (1994)
8. Druzgal, T.J., D’Esposito, M.: Dissecting contributions of prefrontal cortex and fusiform face area to face working memory. *J. Cogn. Neurosci.* **15**, 771–784 (2003)
9. Shucard, J.L., Lee, W.H., Safford, A., Shucard, D.W.: The relationship between processing speed and working memory demand in systemic lupus erythematosus: evidence from a visual n-back task. *Neuropsychology* **24**, 45–52 (2011)
10. McEvoy, L.K., Smith, M.E., Gevins, A.: Dynamic cortical networks of verbal and spatial working memory: effects of memory load and task practice. *Cereb. Cortex* **8**, 563–574 (1998)
11. Watter, S., Geffen, G.M., Geffen, L.B.: The n-back as a duel-talk: P300 morphology under divided attention. *Psychophysiology* **46**, 307–311 (2001)
12. Wintink, A.J., Segalowitz, S.J., Cudmore, L.J.: Task complexity and habituation effects on frontal P300 topography. *Brain Cogn.* **46**, 307–311 (2001)
13. Shucard, J.L., Tekok-Kilic, A., Shiels, K., Shucard, D.W.: Stage and load effects on ERP topography during verbal and spatial working memory. *Brain Res.* **1254**, 49–62 (2009)
14. Bielak, A.A.M., Hultsch, D.F., Strauss, E., McDonald, S.W.S., Hunter, M.A.: Intraindividual variability in reaction time predicts cognitive outcomes five years later. *Neuropsychology* **24**, 731–741 (2010)
15. Mohr, P.N.C., Nagel, I.E.: Variability in brain activity as an individual difference measure in neuroscience. *J. Neurosci.* **30**, 7755–7757 (2010)
16. Specht, C.M., Shucard, D.W.: Single-trial latency variability does not contribute to fast habituation of the long-latency averaged auditory evoked potential in the albino rat. *Electroencephalogr. Clin. Neurophysiol.* **100**, 462–471 (1996)
17. Shucard, D.W., Santa Maria, M.P., Specht, M.C., Podkulski, M.: Single-trial latency variability of auditory evoked potentials may indicate immediate memory in the albino rat. *Int. J. Psychophysiol.* **47**, 229–241 (2003)
18. Usal, A., Segalowitz, S.J.: Sources of P300 attenuation after head injury: single trial amplitude, latency jitter, and EEG power. *Psychophysiology* **32**, 249–256 (1995)
19. Delorme, A., Makeig, S., Fabre-Thorpe, M., Sejnowski, T.: From single trial EEG to brain area dynamics. *Neurocomputing* **44–46**, 1057–1064 (2002)