Single trial variability in neural activity during a working memory task: A window into multiple distinct information processing sequences

Johan Nakuci[1,2*], Thomas J. Covey[1,3], Janet L. Shucard[1,3], David W. Shucard[1,3,4,5], and Sarah F. Muldoon[1,6*]

[1] Neuroscience Program, University at Buffalo, Buffalo, NY
[2] School of Psychology, Georgia Institute of Technology, Atlanta, Georgia, 30332, USA
[3] Department of Neurology, Division of Cognitive and Behavioral Neurosciences, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, Buffalo, NY
[4] Department of Pediatrics, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, Buffalo, NY
[5] Department of Psychology, Jacobs School of Medicine and Biomedical Sciences University at Buffalo, Buffalo, NY
[6] Department of Mathematics and CDSE Program, University at Buffalo, Buffalo, NY

* Johan Nakuci, Sarah F. Muldoon.

Email: jnakuci3@gatech.edu, smuldoon@buffalo.edu

Author Contributions:
J.N. and S.F.M designed the analysis; J.N. and T.J.C processed data; T.J.C, J.L.S. and D.W.S. designed the study and acquired data; J.N., T.J.C, J.L.S., D.W.S and S.F.M interpreted the data and wrote the paper.

Competing Interest Statement: The authors have no competing interests.

Keywords: Working memory, single trial variability, modularity-maximization, event-related potentials

Abstract

Successful encoding, maintenance, and retrieval of information stored in working memory requires persistent coordination of activity among multiple brain regions. It is generally assumed that the pattern of such coordinated activity remains constant for a given task, and therefore to account for noisy data, multiple trials of the same task are performed, and the neural response is averaged across trials to generate an event-related potential (ERP). However, here we show that from trial to trial, the neuronal activity recorded with electroencephalogram (EEG) is actually spatially and temporally diverse, conflicting with the assumption of a single pattern of coordinated response for a given task. We develop a data-driven classification method based on community detection to identify trials with unique spatial and temporal activity and establish that variability in neuronal activity among single time-locked trials arises from the presence of distinct forms of stimulus dependent synchronized activity (i.e., ERPs). The patterns of neuronal activity among trials are limited to a few distinct clusters, or subtypes, that are spatially and temporally distinct and associated with differences in decision-making processes. These results demonstrate that differences in the patterns of neural activity during working memory tasks represent fluctuations in the engagement of distinct brain regions and cognitive processes, suggesting that the brain can choose from multiple mechanisms to perform a given task.

Significance Statement

Working memory is a complex cognitive ability requiring coordinated activity among multiple brain regions to encode, maintain, and retrieve information. It is generally assumed that the pattern coordination between brain regions remains constant and one can average data across multiple trials of the same task. We instead show that there is significant variability in the patterns of brain activity between trials of the same task and develop a method to classify brain activity into distinct subtypes of responses, each with a different spatial and temporal pattern. The subtypes can be

associated with differences in decision-making processes, suggesting that the brain can use multiple mechanisms to perform a given task.

Introduction

Variability is a ubiquitous aspect of real-world, day-to-day activities, cognitive testing, and behavioral performance (1–7). Moment-to-moment neurophysiological variability is present during simple and cognitively demanding tasks, and may impact the efficiency and accuracy in performance (8–11). However, despite the fact that the presence of variability in neuroimaging data has been accepted, it is still often treated as noise during analysis, as opposed to a feature and richness of the data that could lead to insight into brain function.

Working memory tasks provide an arena to study how neurophysiological variability can relate to cognition. Working memory serves as the scaffolding in which we hold and manipulate thoughts and plan actions. As such, it is a cognitively demanding process requiring coordinated activity among multiple brain regions to encode, maintain, and retrieve information (12, 13). Despite the need to maintain this coordinated activity, variability in neuronal activity across trials is pervasive during a working memory task (6, 14). Such fluctuations are typically considered noise and eliminated through averaging neuronal activity across trials to identify stimulus dependent synchronized activity (15–18).

In electroencephalogram (EEG) recordings, this stimulus dependent synchronized activity can be identified with event-related potentials (ERPs) (Luck, 2014)which provide the average time-locked response to a stimulus locked trial (19). However, averaging trials may obscure important information (19) since trial-to-trial frequency and amplitude differences can arise from (a) fluctuations in postsynaptic response (20), (b) alterations in the balance between excitatory and inhibitory neural systems (21), (c) competition between brain networks (22), and (d) changes in attention (23) as discussed further by Dinstein *et. al* 2015. Additionally, averaging might eliminate the ability to observe distinct forms of synchronized activity (i.e., the ERP) that are engaged differentially across individual trials or contexts, but that are, nonetheless, relevant to behavioral performance. This issue may be particularly true for tasks with distinct types of trials, or when

cognitive processing engages multiple subprocesses that rely on complex interactions between distinct neural networks (24). Therefore, distinct types of synchronized activity could indicate changes in complex interactions or operations among distributed neural networks during working memory (25).

In the present study, we investigate variability in neural activity across individual trials to determine whether the variability is indicative of distinct forms of stimulus dependent synchronized activity important for working memory and behavioral performance. We use an *n*-back working memory task with three levels of difficulty, 0-back (N_{trials} = 3150, N_{subj} = 21), 1-back (N_{trials} = 3108, N_{subj} = 21), and 2-back (N_{trials} = 3066, N_{subj} = 21), to characterize trial-to-trial variability in neural activity (Fig. 1A). We develop a data-driven clustering method to classify individual trials based on differences in spatial and temporal EEG activity recorded from the scalp. This approach leverages the spatial-temporal resolution of dense-electrode EEG recording to identify distinct characteristics of neuronal activity.

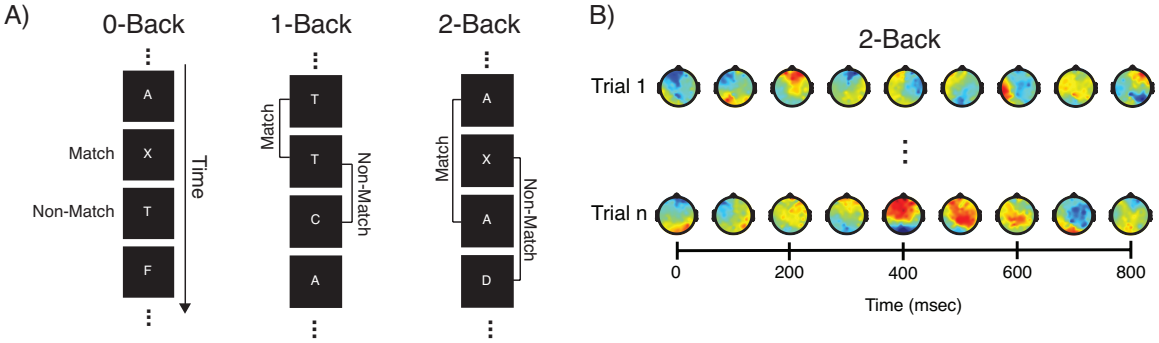

Figure 1. Task description and trial-to-trial variability in neuronal activity. A) An example of the visual 0-, 1- and 2-back task. For the 0-back condition, a match trial was one in which the letter presented was an “X” and the non-match for all other letters. For the 1-back condition, a match trial is one in which the second letter “T” matches the letter one trial before. Similarly, a non-match trial is one in which the letter “C” does not match the letter presented one trials back, “T”. In the 2-back condition, a match trial is one in which the second letter “A” matches the letter presented two trials ago. Similarly, a non-match trial is one in which the letter “D” does not match the letter presented two trials back, “X”. B) EEG activity from two trials part of the 2-back condition from stimulus onset (0 msec) to 800 msec from the same subject. The brain activity between the trials exhibits stark differences particularly after 200 msec.

Results

During the *n*-back working memory task, trial-to-trial variability is marked by the presence of distinct spatial and temporal patterning in neuronal activity (Fig. 1B). We investigated if these differences in neural activity indicated the existence of distinct forms of stimulus dependent synchronized activity. We developed a data-driven clustering method based on community detection that utilizes topographical and temporal differences to classify individual trials. The pairwise similarity in neuronal activity across trials was determined using the Pearson Product Correlation of full brain EEG sensor activity from stimulus onset (0 msec) to 852 msec (144 sensors x 750 time points per sensor). This trial-by-trial similarity matrix was then clustered using Modularity-Maximization, an unsupervised clustering method that does not require the number of clusters to be specified (26). As shown in Fig. 2A for the 2-back task, modularity-maximization separated individual trials into 3 clusters, or subtypes we label T1 (N_{trials} = 1134), T2 (N_{trials} = 1061), and T3 (N_{trials} = 871) (Fig. 2A).

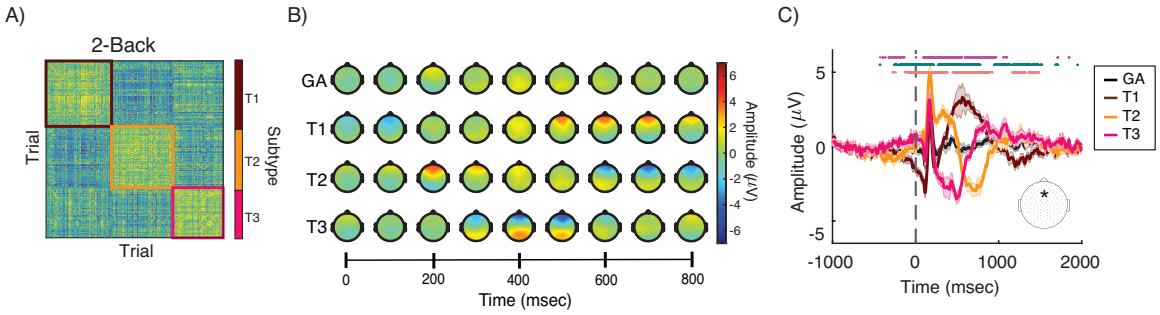

Figure 2. Subtypes of Individual Trials in the 2-back. A) Modularity-Maximization based clustering identified 3 subtypes of trials, T1, T2, and T3. The colored squares correspond to the trials composing each subtype. Pearson Correlation was used to calculate the spatial temporal similarity of the EEG activity from individual trial. B) Spatial and temporal ERP activity from the T1, T2, T3 subtypes and the grand average (GA) at 9 time points post-stimulus, respectively. C) ERP activity from frontal-midline electrode. Each waveform shows the mean (thick line) and standard error of the mean (shaded area). Insert shows the location of the EEG sensor from which the ERP waveforms were derived. Statistical testing is based on Paired T-test and FDR corrected for multiple comparison (*$p_{corrected}$ < 0.05). Statistically significant differences in amplitude are marked at the top of each panel (the three rows of colored asterisks) between: * T1 and T2; * T1 and T3; * T2 and T3.

ERPs are neuronal activity derived from spontaneous and synchronized activity generated in response to a stimulus and averaged across trials for a time-locked task (19). To better assess the potential novelty of the analytical approach used here, we examined the correspondence between the spatial and temporal properties of the conventional signal-averaged ERP waveform (the grand average (GA) generated by averaging across all trials regardless of trial type) and the ERPs acquired by separately averaging across trials within each subtype (T1, T2, T3) obtained from the modularity-maximization analysis (Fig. 2B-C). Qualitatively, the topography of the ERPs from each subtype in the 2-back condition fluctuated between frontal and parietal positive amplitude relative to each other and the grand average (Fig. 2B; SI Movie 1). Further, these topographical fluctuations show significantly different temporal profiles among each other when focusing on a single EEG sensor (Fig. 2C).

Expanding this analysis to the 0- and 1-back condition, similarly, modularity-maximization identified 3 subtypes for the 0-back [T1 (N_{trials} = 951), T2 (N_{trials} = 1101), and T3 (N_{trials} = 1098) and 1-back [T1 (N_{trials} = 899), T2 (N_{trials} = 987), and T3 (N_{trials} = 1222)] (Fig. 3A and D, respectively). The topography of these ERPs fluctuated between frontal and parietal positivity (Fig. 3B and E, SI Movie 2 & 3, respectively) and with significant amplitude differences from a single sensor, reflecting the findings from the 2-back (Fig. 3C and F, respectively). Notably, each of these distinct waveforms was observed for different trials within a given individual indicating that the subtypes are not simply due to differences between individuals (SI Fig. 1).

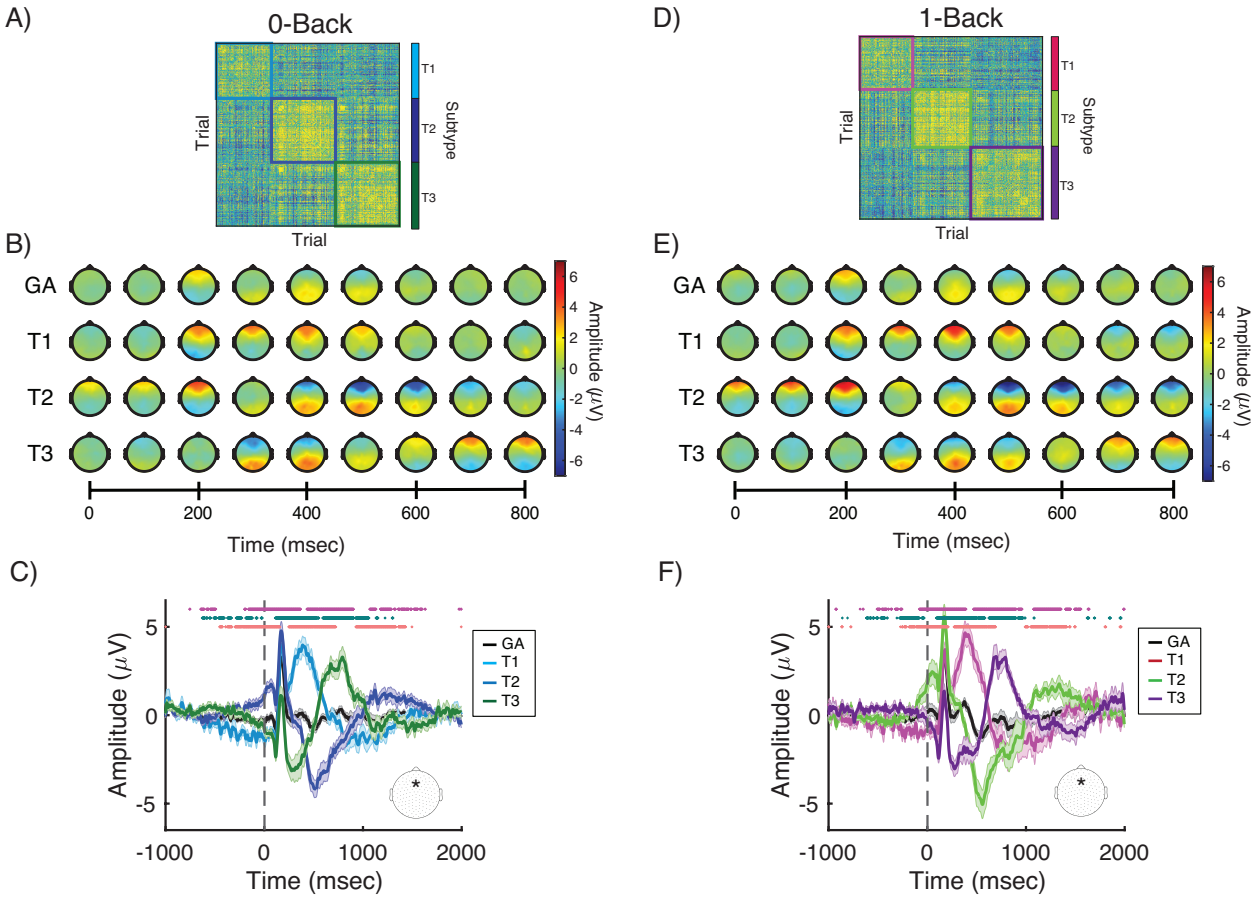

Figure 3. Subtypes of Individual Trials in the 0 and 1-back. A and D) Modularity-Maximization based clustering identified 3 subtypes of trials, T1, T2, and T3 in 0- and 1-back, respectively. The colored squares correspond to the trials composing each subtype. B and E) Spatial and temporal ERP activity from the T1, T2, T3 subtypes and the grand average (GA) at 9 time points post-stimulus, respectively. C and F) ERP activity from frontal-midline electrode. Each waveform shows the mean (thick line) and standard error of the mean (shaded area). Insert shows the location of the EEG sensor from which the ERP waveforms were derived. Statistical testing is based on Paired T-test and FDR corrected for multiple comparison (*$p_{corrected} < 0.05$). Statistically significant differences in amplitude are marked at the top of each panel (the three rows of colored asterisks) between: * T1 and T2; * T1 and T3; * T2 and T3.

We quantified the spatial-temporal similarity between ERPs from subtype to traditional grand-average to determine if subtypes contain ERPs different from the average. The spatial-temporal similarity was estimated using Pearson Correlation from 0msec to 852msec. The grand-average ERP was strongly ($r > 0.75$) correlated with only one of the ERP subtypes in the 0-back (SI Fig. 2A, *Middle*) and 1-back (SI Fig. 2B, *Middle*), and moderately ($r > 0.50$) correlated with one subtype in the 2-back (SI Fig. 2C, *Bottom*). This supports work suggesting the traditional grand

average ERP, in the best case, captures only a portion of true stimulus-driven synchronized brain activity and in the worst case, the grand average ERP obscures the actual brain activity associated with cognitive processing (19, 27).

The above analysis gives only an overarching similarity, we therefore focused on identifying the specific time points the ERPs differ. We quantified the similarity at each time point between subtypes and to the grand average. For all ERP subtypes, there was a period of high topographical similarity (r > 0.75) between approximately 152msec to 200msec (Fig. 4A-C). This period is associated with visual processing and attention indicating that the divergence among the subtype waveforms is not due to differences in early visual processing or differences in attention at stimulus onset (28).

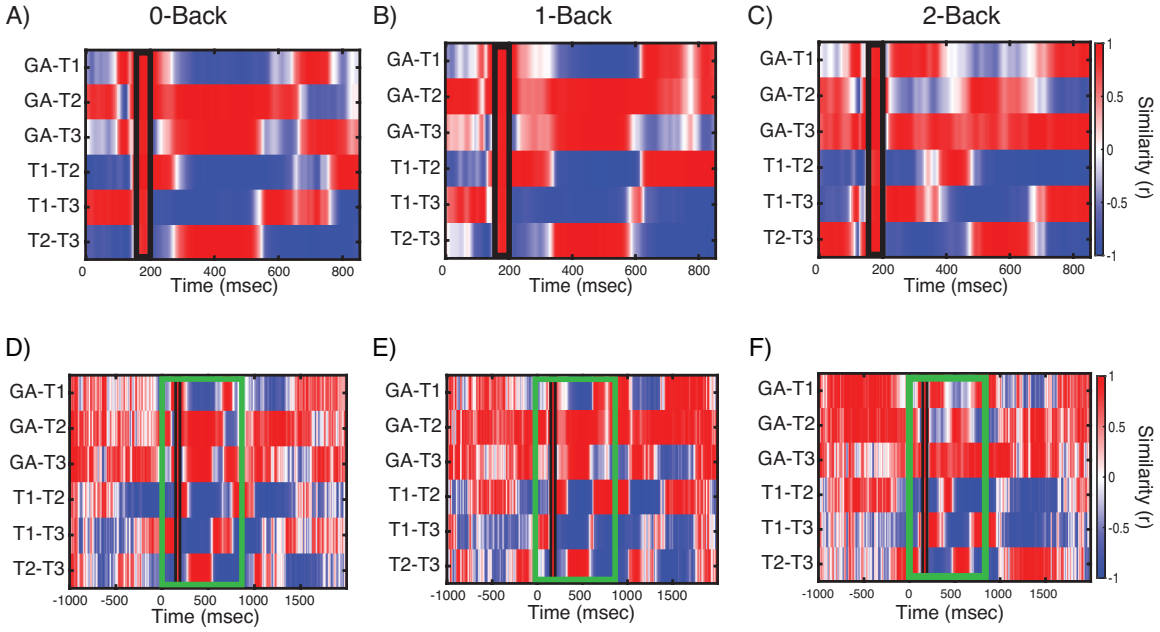

Figure 4. Spatial and temporal heterogeneity in ERP active from each of the subtypes. A-C) The similarity in spatial topography using the Pearson Product Correlation among the ERP subtypes and the grand average for the 0-, 1- and 2-back, respectively. The spatial similarity flips between positively and negatively correlated with a brief period between with high spatial similarity among the GA and subtypes (black box). The spatial similarity was calculated at each time point between stimulus onset (0msec) to 852 msec. D-F) The spatial similarity from -1000 to 2000msec post-stimulus. In panels D-F, the green box corresponds to the period in panel A-C.

Outside of this approximately 50msec window, the ERP waveforms fluctuated between positively to negatively correlated topographies suggesting the different ERP waveforms are the result of brain regions/networks becoming active at different periods in the trial. These topographical fluctuations extended to the pre-stimulus and after 852msec, indicating they were not an effect of the of the time range used for classification (Fig. 4D-F). The presence of these fluctuations in the pre-stimulus time range suggests that pre-stimulus activity might play a role in determining post-stimulus activity (29) and behavioral performance (30–33).

The existence of these subtypes raised the possibility that they were relevant for behavioral performance. We tested for differences in reaction times and the proportion of correct responses for match/non-match trials among subtypes. We found significant differences in reaction times between subtypes in the 0-back (T1 = 558 $\pm$ 209.75; T2 = 513 $\pm$ 219; T3 = 506 $\pm$ 209.25 [Median $\pm$ IQR msec]; Kruskal-Wallis Test, $\chi^2 = 43.56$, $p = 3.47 \times 10^{-10}$), 1-back (T1 = 699 $\pm$ 261.75; T2 = 624 $\pm$ 257.50; T3 = 620 $\pm$ 302 [Median $\pm$ IQR msec]; Kruskal-Wallis Test, $\chi^2 = 92.51$, $p = 8.14 \times 10^{-21}$), and 2-back (T1 = 767.50 $\pm$ 374; T2 = 747 $\pm$ 325.25; T3 = 705 $\pm$ 368.5 [Median $\pm$ IQR msec]; Kruskal-Wallis Test, $\chi^2 = 18.8$, $p = 7.95 \times 10^{-5}$) suggesting subtypes were associated with decision making processes (Fig. 5A-C). However, these processes were not associated with stimulus categorization because the subtypes contained the same proportion of correct match/non-match trials (ANOVA, $p > 0.05$; Fig. 5D-F). Thus, not only do the subtypes appear to have ERPs with spatial-temporal activity that differs from each other and from the traditional grand-average, they also are associated with differences in behavioral performance.

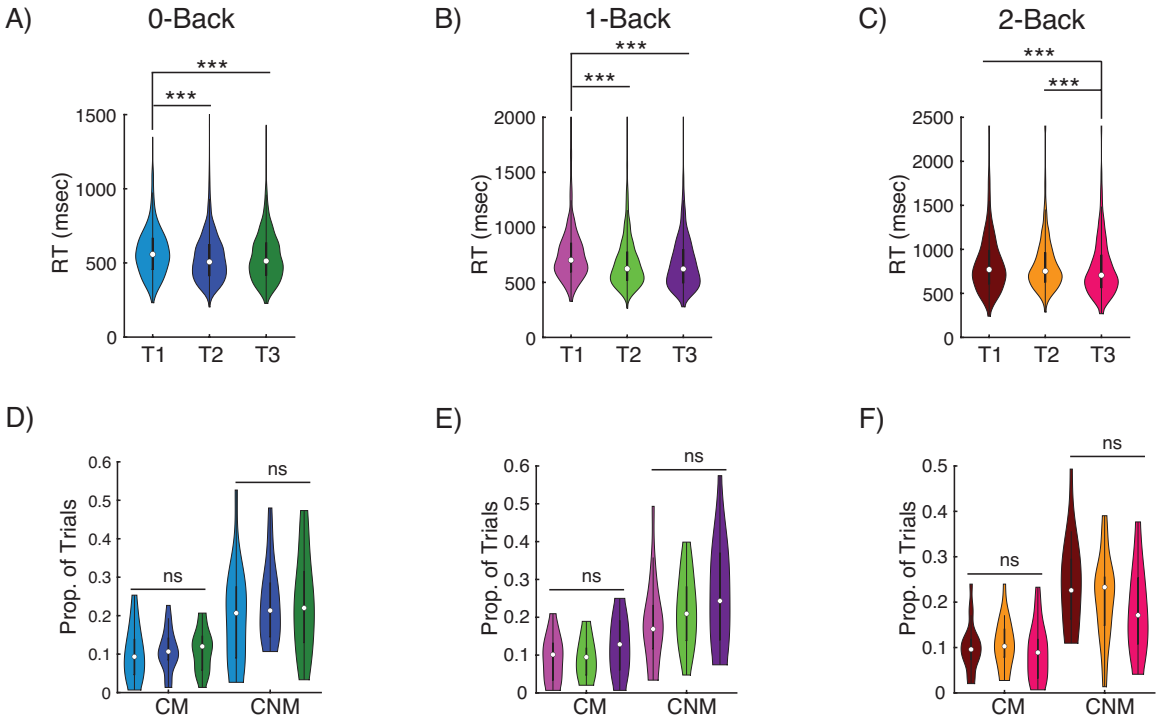

Figure 5. Behavioral Associations Among Trial Subtypes. A-C) Differences in reaction times between trial subtypes for the 0-, 1-, 2-back, respectively. Post-hoc statistical tests were performed with Wilcoxon rank sum test and Bonferroni multiple comparison correction. D-F) The proportion of correct match (CM) and correct non-match (CNM) responses was equivalent between trial subtypes for the 0-, 1-; and 2-back, respectively. Violin plot shows the distribution of the data with the circle representing the median value, box containing the 25th and 75th percentiles and line representing all data that are not considered outliers. *** $p_{corrected} < 0.001$, ns = not significant.

We next investigated if the spatial temporal activity of the ERP subtypes were similar across the *n*-back conditions. In each condition, subgroups are labeled T1, T2, and T3 in decreasing order of the reaction time associated with the subtype (i.e., T1 is always the subtype associated with the slowest reaction time). As seen in Fig. 6A, when comparing the spatial-temporal profiles of each subtype with the corresponding subtype (i.e., T1 vs T1, T2 vs. T2, T3 vs. T3), we found strong ($r > 0.75$) similarity in spatio-tempotal patterning of subtypes between 0- and 1-back tasks. When comparing the 2-back task with both the 0- and 1-back tasks (Fig. 6B-C), we still observe moderate ($r > 0.50$) similarity between corresponding T2 and T3 subtypes, although the similarity for T1 subtypes weakens significantly. Interestingly, we observe the lowest values of similarity

when comparing 0- and 2-back tasks, suggesting that the difficulty of the task modifies the spatio-temporal profile of the subtyped ERPs.

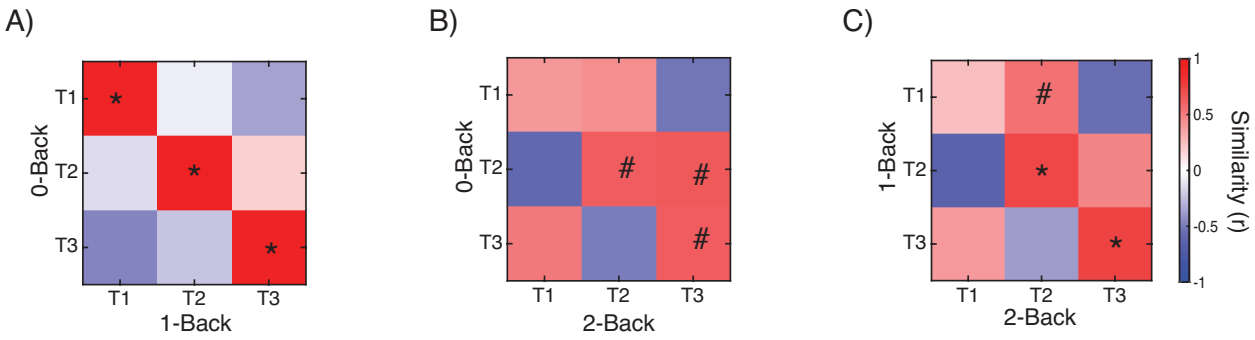

Figure 6. Similarity of ERP subtypes between *n*-back conditions. Spatial-temporal ERP similarity among subtypes between the A) 0- and 1-back condition; B) 0- and 2-back condition; and C) 1- and 2-back condition. The spatial-temporal similarity was estimated using Pearson Correlation from 0msec to 852msec. * r > 0.75; # r > 0.50. $p < 10^{-15}$

Discussion

Trial-to-trial variability is ubiquitous in neurophysiological activity during simple and cognitively demanding tasks (34, 35). This raises questions about the role neuronal variability plays in brain function and cognition. Variability in neuronal activity across different stimulus categories may provide a benefit in that it allows flexibility when responding to changing demands (36). However, this variability may need to be calibrated appropriately to produce a beneficial response because otherwise it could be detrimental to neural performance (37–39) and/or may reflect underlying functional/structural disturbances (40–43). In the present study, we investigated if variability in neural activity indicates the existence of distinct forms of stimulus dependent synchronized activity among trials reflecting different cognitive processes.

We developed a data-driven topographical method to classify neuronal activity measured with EEG from individual trials during a visual *n*-back task, with three level of difficulty: 0-, 1- and 2-back. This analysis allowed us to define not only a standard GA ERP, but also ERPs for each of

the subtypes identified from the clustering analysis, in order to investigate the relation between stimulus driven synchronized activity with working memory and behavioral performance. We limited our analysis primarily to ERPs because the ERP is a well characterized form of synchronized brain activity generated in response to a stimulus (19).

We found that trial-to-trial variability in neuronal activity was limited to a few distinct waveforms or subtypes and these subtypes contained ERPs with unique spatial and temporal organization. The distinct subtypes were associated with differences in decision-making processes. Furthermore, traditional averaging, in the best case, captures only a portion of true stimulus-driven synchronized brain activity and in the worst case, traditional averaging obscured the actual brain activity associated with cognitive processes (19, 27).

Spatial and temporal differences among subtype ERP waveforms were not due to differences in the early visual processing and attention since the period between 152-200msec was marked by high topographical similarity. This period of similarity in neuronal activity is in line with previous work using linear discriminant analysis to classify individual-trials in which the early EEG component was not associated with a participant's ability to discriminate between choices (44, 45). Therefore, the divergence among the subtypes/ERP waveforms was not due to differences in the early visual processes or differences in attention, but likely due later stage cognitive processes.

Outside of this 50 msec period, the neurophysiological activity fluctuated between positively- and negatively correlated ERP topographies. For example, the topographies of the T1 and T2 subtypes in the 2-back were negatively correlated from approximately 200 to 300 msec and positively correlated between 400 to 500 msec post-stimulus. This flickering could be a marker of distinct brain regions and cognitive processes being recruited to accomplish the task. Engagement of different cognitive processes and their corresponding brain regions could be a

result of having to make a decision under varying levels of uncertainty or conflict between choices (46). In the *n*-back task, uncertainty and conflict could arise from the maintenance and/or decision component because, unlike other working memory tasks, the *n*-back integrates these two components (47).

Furthermore, these unique forms of stimulus-driven brain activity (subtypes) effected behavioral performance. Trials with slowed response speed were associated with ERP subtypes that contained a positive frontal component at approximately 300 msec, which is aligned with the canonical spatial-temporal characteristics of the novelty-P3 (48) and with enhancement of executive function associated with increased task difficulty (49). The novelty-P3 is associated with equivocation during cognitive performance suggesting participants may be perceiving a stimulus that they did not expect (48, 50). On the other hand, efficient (faster) performance was associated with frontal negativity at ~250msec, the N2-P3 complex (51), and followed by frontal positivity at ~400msec, the N400-late positivity complex (52). The combination of neurophysiological and behavioral effects supports the conclusion that distinct brain regions and cognitive processes are recruited to accomplish the task.

Interestingly, we did observe a strong similarity in the spatio-temporal profiles of ERP subtypes between the 0- and 1-back tasks. While further work should investigate the exact cognitive processes involved in each subtype, this suggests that the brain switches between similar cognitive processes for each of these tasks. The decrease in similarity when comparing subtasks from the 0- or 1-back task with the 2-back task suggests that task difficulty might also modify the exact cognitive processes involved. It is also interesting to observe that the spatio-temporal profile of the T2 and T3 subtypes are more similar between the 0- and 2-back tasks, whereas the T1 and T2 subtypes show increased similarity between the 1- and 2-back tasks. Future work should investigate more specifically how task difficulty modifies these spatio-temporal profiles and how this relates to the cognitive processes involved in performing the task.

Lastly, although the analysis was limited to EEG and ERPs; our methodological and analytical approach can be applied to other high-dimensional neurophysiological recordings such as individual neurons (53). Additionally, this analytical approach also can be applied to activity from frequency bands obtained during task-induced EEG activity, another widely used electrophysiological method for evaluating cognitive function (54).

In summary, these results indicate that working memory is dependent on a set of limited, but diverse sequence of cognitive processes reflecting coordinated neuronal activity from spatially and temporally diverse regions. Why the brain might benefit from having multiple but distinct processes to perform the same function remains an open question. One might speculate that it could relate to robustness of task performance or depleted resources for a given brain region over time. Regardless, our results indicate that in order to answer these questions, it will be necessary to perform a more nuanced analysis of neuroimaging data during task performance, as valuable information is lost by averaging across trials. This loss can be mitigated through classifying individual trials to estimate the "real" stimulus-driven synchronized neuronal activity contributing to working memory and cognitive function, and we encourage future work to embrace this approach.

Materials and Methods

Participants and task

A total of 41 healthy participants completed a visual *n-back* task. Participants were recruited as part of a larger study examining cognitive disturbances related to Systemic Lupus Erythematosus The current analyses were based on 21participants (17 Female; Age = 42 $\pm$ 12 years) who completed a visual *n-back* task (0-, 1-, 2-back). The number of participants in the current analysis was lower than that of the total number in the study because participants were removed if more than 5% of the trials from any of the n-back tasks exceeded 100 μV which likely was due to

artifacts such as eye-blinks and movement. The removal of participants instead of trials, ensured that all participants had the same number of trials for analysis and contributed equally to the clustering analysis. All participants were required to have at least 50 percent correct trials to be included in the study.

All participants analyzed in the present study were neurologically normal. Participants were excluded if they had history of head trauma, visual problems that could not be corrected, learning disability, neurological impairment, or if they had an Axis I psychiatric disorder. Participants provided informed consent and received financial compensation for participating in the study. This research was approved by the University at Buffalo Institutional Review Board.

Each participant completed a visual n-back task (0-, 1-, 2-back) with a stimulus duration of 400ms, and 2000ms inter-stimulus-interval. Letters were presented at the center of a computer monitor, approximately 70 cm from the participant's nasion. The *n-back* conditions were administered in the following order: 0-back, 1-back, and 2-back. A practice trial preceded each condition. In each of the conditions a total of 150 letters were presented one at a time on a computer screen via E-Prime software (E-Prime, 2002). Each condition was divided into two blocks of 75 letters. A brief break was taken between blocks and an approximately 5-minute break was taken between conditions. During the task, participants were instructed to press the inner buttons of a four -button response pad, simultaneously with both thumbs, if the letter they observed on the screen matched the letter presented *n* trials back. They pressed the outer buttons of the response pad with both thumbs if the letter presented on the screen did not match the letter presented *n* trials back. For the 0-back condition, participants were instructed to press the "match" buttons if an "X" was presented, and the "non-match" buttons for all other letters. Fifty of the letters in each condition were match trials and 100, 98, and 96 were a non-match trials in the 0-, 1-, and 2-Back conditions, respectively (the first letter of each block for the 1-back, and the first two letters of each block for the 2-back were not counted due to the task requirements.

Reaction time (RT) to both match and nonmatch trials, correct responses, and EEG were simultaneously recorded for each condition.

EEG recording and processing

Participants were seated in a comfortable chair in a dimly lit, sound attenuated room in front of a computer monitor. A 256-channel dense electrode HydroCel Geodesic Sensor Net (Phillips-Electrical Geodesics, Inc., Eugene, OR) was fitted onto the participant's head. During EEG acquisition, electrode impedances were kept below 50 kΩ whenever possible. A vertex reference was used during recording, and EEG data were filtered between 0.1 to 100 Hz and digitized at a sampling rate of 250 Hz.

The EEG cleaning procedure was implemented using EEGLAB v14 (55). EEG data were band-pass filtered between 0.5 Hz and 50 Hz using a Finite Impulse Response filter. Next, artifact subspace reconstruction (ASR) was used to determine and remove bad channels (56). Following ASR, independent component analysis (ICA) decomposition, and ADJUST was used to remove ICA components associated with noise (eye blinks, muscle artefact, etc.) (57). Following ICA decomposition, bad channels were interpolated using spherical interpolation.

ASR was used to interpolated data segments not cleaned by ICA. Data was scanned with a 500msec sliding window to find segments that are greater than four standard deviations from the reference segment. These noisy segments are treated as missing data and reconstructed. Lastly, the data were re-referenced to a common average reference and were segmented from -1000 msec to 2000 msec post-stimulus, and baseline corrected by subtracting the mean amplitude of the baseline portion (-1000 to 0 msec) of the epoch.

Modularity-maximization clustering analysis

Prior to clustering, the first two trials in the 2-back were removed because they were non-response trials. Next, data were pooled from all 21 participants resulting in 3066 trials for the 2-back. A trial-by-trial similarity matrix was created using the Pearson Product Correlation between trials calculated from 0 msec (stimulus onset) to 852 msec across 144 sensors above the canthomeatal line, corresponding to 30,816 data points per trial. Clustering was limited to an interval from 0 msec to 852 msec post-stimulus because: 1) this time interval contains stimulus driven activity; 2) the 852 msec time occurs after the median reaction time of 745msec, ensuring the relevant brain activity related to the trial that leads to the eventual behavioral response is maintained; and 3) this minimized interference from EEG artifacts (i.e. eye blinks) that tended to occur most frequently in the inter-stimulus-interval (after 852 msec).

Clustering of the similarity matrix was conducted using modularity-maximization (26). Modularity-maximization identifies clusters in which the similarity within a cluster is greater than between clusters. In addition, modularity-maximization does not require the number of clusters to be specified. The resolution of the cluster can be controlled with resolution parameter; here we use $\gamma = 1$. Modularity-maximization was implemented with the Generalized Louvain algorithm part of the Brain Connectivity Toolbox (58). Clustering results from the Generalized Louvain algorithm can be dependent on the initial seeding. To avoid such suboptimal results, clustering was repeated 100x and final clustering results were based on the consensus across iterations (59).

Identification of event-related potentials (ERPs)

ERPs for each subtype were generated by averaging all trials within each of the corresponding clusters. In addition to ERP Subtypes, grand averaged ERPs were generated to assess the similarity in topography obtained between a standard approach to ERP analysis and that of the clustering methodology outlined above. Grand averaged ERPs were generated from averaging all trials collapsed in the 2-back task.

Statistical analysis and software

Statistical testing used both parametric and non-parametric tests. To test for ERP amplitude differences, we used Paired T-test and False Discovery Rate (FDR) multiple comparison correction. Statistical differences for reaction times were tested using Kruskal-Wallis test and post-hoc comparison was conducted using Wilcoxon rank sum test with Bonferroni multiple comparison correction because of the non-parametric distribution of the corresponding values. Whereas differences in accuracy were tested with ANOVA. All data processing and statistical analyses were performed in MATLAB 2018b.

Data Sharing

All data and code is available via request to the authors upon publication.

Acknowledgments

This work was supported by National Institutes of Health Grant NS049111, PI JLS.

References

1. G. J. Tomko, D. R. Crapper, Neuronal variability: non-stationary responses to identical visual stimuli. *Brain Research* (1974) https:/doi.org/10.1016/0006-8993(74)90438-7.

2. R. L. T. Goris, J. A. Movshon, E. P. Simoncelli, Partitioning neuronal variability. *Nature Neuroscience* (2014) https:/doi.org/10.1038/nn.3711.

3. A. Arieli, A. Sterkin, A. Grinvald, A. Aertsen, Dynamics of ongoing activity: Explanation of the large variability in evoked cortical responses. *Science (1979)* (1996) https:/doi.org/10.1126/science.273.5283.1868.

4. M. N. Shadlen, W. T. Newsome, The variable discharge of cortical neurons: Implications for connectivity, computation, and information coding. *Journal of Neuroscience* (1998) https:/doi.org/10.1523/jneurosci.18-10-03870.1998.

5. M. Carandini, Amplification of trial-to-trial response variability by neurons in visual cortex. *PLoS Biology* (2004) https:/doi.org/10.1371/journal.pbio.0020264.

6. M. Shafi, *et al.*, Variability in neuronal activity in primate cortex during working memory tasks. *Neuroscience* (2007) https:/doi.org/10.1016/j.neuroscience.2006.12.072.

7. B. Blankertz, S. Lemm, M. Treder, S. Haufe, K. R. Müller, Single-trial analysis and classification of ERP components - A tutorial. *Neuroimage* (2011) https:/doi.org/10.1016/j.neuroimage.2010.06.048.

8. L. Waschke, M. Wöstmann, J. Obleser, States and traits of neural irregularity in the age-varying human brain. *Scientific Reports* **7**, 17381 (2017).

9. A. R. McIntosh, N. Kovacevic, R. J. Itier, Increased Brain Signal Variability Accompanies Lower Behavioral Variability in Development. *PLOS Computational Biology* **4**, e1000106- (2008).

10. B. Mišić, T. Mills, M. J. Taylor, A. R. McIntosh, Brain Noise Is Task Dependent and Region Specific. *Journal of Neurophysiology* **104**, 2667–2676 (2010).

11. R. M. G. Reinhart, J. A. Nguyen, Working memory revived in older adults by synchronizing rhythmic brain circuits. *Nature Neuroscience* **22**, 820–827 (2019).

12. W. J. Ma, M. Husain, P. M. Bays, Changing concepts of working memory. *Nature Neuroscience* (2014) https:/doi.org/10.1038/nn.3655.

13. M. D'Esposito, B. R. Postle, The Cognitive Neuroscience of Working Memory. *Annual Review of Psychology* (2015) https:/doi.org/10.1146/annurev-psych-010814-015031.

14. D. W. Shucard, T. J. Covey, J. L. Shucard, Single Trial Variability of Event-Related Brain Potentials as an Index of Neural Efficiency During Working Memory in *Foundations of Augmented Cognition: Neuroergonomics and Operational Neuroscience*, D. D. Schmorrow, C. M. Fidopiastis, Eds. (Springer International Publishing, 2016), pp. 273–283.

15. L. Busse, A. R. Wade, M. Carandini, Representation of Concurrent Stimuli by Population Activity in Visual Cortex. *Neuron* **64**, 931–942 (2009).

16. P. Bao, L. She, M. McGill, D. Y. Tsao, A map of object space in primate inferotemporal cortex. *Nature* **583**, 103–108 (2020).

17. I. Nauhaus, A. Benucci, M. Carandini, D. L. Ringach, Neuronal Selectivity and Local Map Structure in Visual Cortex. *Neuron* **57**, 673–679 (2008).

18. G. B. Keller, T. Bonhoeffer, M. Hübener, Sensorimotor Mismatch Signals in Primary Visual Cortex of the Behaving Mouse. *Neuron* **74**, 809–815 (2012).

19. S. J. (Steven J. Luck, *An Introduction to the Event-Related Potential Technique, second edition* (2014) https:/doi.org/10.1118/1.4736938.

20. C. Ribrault, K. Sekimoto, A. Triller, From the stochasticity of molecular processes to the variability of synaptic transmission. *Nature Reviews Neuroscience* (2011) https:/doi.org/10.1038/nrn3025.

21. C. van Vreeswijk, H. Sompolinsky, Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science (1979)* (1996) https:/doi.org/10.1126/science.274.5293.1724.

22. A. M. Clare Kelly, L. Q. Uddin, B. B. Biswal, F. X. Castellanos, M. P. Milham, Competition between functional brain networks mediates behavioral variability. *Neuroimage* (2008) https:/doi.org/10.1016/j.neuroimage.2007.08.008.

23. A. Fontanini, D. B. Katz, Behavioral states, network states, and sensory response variability. *Journal of Neurophysiology* (2008) https:/doi.org/10.1152/jn.90592.2008.

24. Y.-N. Chen, S. Mitra, F. Schlaghecken, Sub-processes of working memory in the N-back task: An investigation using ERPs. *Clinical Neurophysiology* **119**, 1546–1559 (2008).

25. A. Baddeley, Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology* (2012) https:/doi.org/10.1146/annurev-psych-120710-100422.

26. P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, J. P. Onnela, Community structure in time-dependent, multiscale, and multiplex networks. *Science (1979)* (2010) https:/doi.org/10.1126/science.1184819.

27. A. Mouraux, G. D. Iannetti, Across-trial averaging of event-related EEG responses and beyond. *Magnetic Resonance Imaging* (2008) https:/doi.org/10.1016/j.mri.2008.01.011.

28. S. J. Luck, E. S. Kappenman, "ERP Components and Selective Attention" in *The Oxford Handbook of Event-Related Potential Components*, (2012) https:/doi.org/10.1093/oxfordhb/9780195374148.013.0144.

29. G. Thut, C. Miniussi, J. Gross, The Functional Importance of Rhythmic Activity in the Brain. *Current Biology* **22**, R658–R663 (2012).

30. A. Bompas, P. Sumner, S. D. Muthumumaraswamy, K. D. Singh, I. D. Gilchrist, The contribution of pre-stimulus neural oscillatory activity to spontaneous response time variability. *Neuroimage* **107**, 34–45 (2015).

31. J. Britz, C. M. Michel, Errors can be related to pre-stimulus differences in ERP topography and their concomitant sources. *Neuroimage* **49**, 2774–2782 (2010).

32. S. Hanslmayr, *et al.*, Prestimulus oscillations predict visual perception performance between and within subjects. *Neuroimage* **37**, 1465–1473 (2007).

33. R. Rajagovindan, M. Ding, From Prestimulus Alpha Oscillation to Visual-evoked Response: An Inverted-U Function and Its Attentional Modulation. *Journal of Cognitive Neuroscience* **23**, 1379–1394 (2011).

34. D. D. Garrett, *et al.*, Moment-to-moment brain signal variability: A next frontier in human brain mapping? *Neuroscience & Biobehavioral Reviews* **37**, 610–624 (2013).

35. A. A. Faisal, L. P. J. Selen, D. M. Wolpert, Noise in the nervous system. *Nature Reviews Neuroscience* **9**, 292–303 (2008).

36. L. Waschke, N. A. Kloosterman, J. Obleser, D. D. Garrett, Behavior needs neural variability. *Neuron* **109**, 751–766 (2021).

37. D. D. Garrett, N. Kovacevic, A. R. McIntosh, C. L. Grady, The Importance of Being Variable. *The Journal of Neuroscience* **31**, 4496 (2011).

38. S. Dave, T. A. Brothers, T. Y. Swaab, 1/f neural noise and electrophysiological indices of contextual prediction in aging. *Brain Research* **1691**, 34–43 (2018).

39. C. L. Grady, D. D. Garrett, Brain signal variability is modulated as a function of internal and external demand in younger and older adults. *Neuroimage* **169**, 510–523 (2018).

40. G. J. Yang, *et al.*, Altered global brain signal in schizophrenia. *Proceedings of the National Academy of Sciences* **111**, 7438 (2014).

41. J. S. Nomi, T. S. Bolt, C. E. C. Ezie, L. Q. Uddin, A. S. Heller, Moment-to-Moment BOLD Signal Variability Reflects Regional Changes in Neural Flexibility across the Lifespan. *The Journal of Neuroscience* **37**, 5539 (2017).

42. D. D. Garrett, *et al.*, Moment-to-moment brain signal variability: A next frontier in human brain mapping? *Neuroscience & Biobehavioral Reviews* **37**, 610–624 (2013).

43. I. Dinstein, D. J. Heeger, M. Behrmann, Neural variability: Friend or foe? *Trends in Cognitive Sciences* (2015) https:/doi.org/10.1016/j.tics.2015.04.005.

44. M. G. Philiastides, P. Sajda, Temporal Characterization of the Neural Correlates of Perceptual Decision Making in the Human Brain. *Cerebral Cortex* **16**, 509–518 (2006).

45. R. Ratcliff, M. G. Philiastides, P. Sajda, Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. *Proceedings of the National Academy of Sciences* **106**, 6539 (2009).

46. M. M. Botvinick, Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive, Affective and Behavioral Neuroscience* (2007) https:/doi.org/10.3758/CABN.7.4.356.

47. S. Watter, G. M. Geffen, L. B. Geffen, The n-back as a dual-task: P300 morphology under divided attention. *Psychophysiology* (2001) https:/doi.org/10.1111/1469-8986.3860998.

48. J. Polich, Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology* **118**, 2128–2148 (2007).

49. S. Dong, L. M. Reder, Y. Yao, Y. Liu, F. Chen, Individual differences in working memory capacity are reflected in different ERP and EEG patterns to task difficulty. *Brain Research* **1616**, 146–156 (2015).

50. H. J. van der Horn, *et al.*, Altered Wiring of the Human Structural Connectome in Adults with Mild Traumatic Brain Injury. *Journal of Neurotrauma* **34**, 1035–1044 (2017).

51. P. D. Gajewski, P. Stoerig, M. Falkenstein, ERP—Correlates of response selection in a response conflict paradigm. *Brain Research* **1189**, 127–134 (2008).

52. E. F. Lau, C. Phillips, D. Poeppel, A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience* **9**, 920–933 (2008).

53. J. J. Jun, *et al.*, Fully integrated silicon probes for high-density recording of neural activity. *Nature* **551**, 232–236 (2017).

54. J. Onton, A. Delorme, S. Makeig, Frontal midline EEG dynamics during working memory. *Neuroimage* **27**, 341–356 (2005).

55. A. Delorme, S. Makeig, EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* (2004) https:/doi.org/10.1016/j.jneumeth.2003.10.009.

56. T. Mullen, *et al.*, Real-time modeling and 3D visualization of source dynamics and connectivity using wearable EEG in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, (2013) https:/doi.org/10.1109/EMBC.2013.6609968.

57. A. Mognon, J. Jovicich, L. Bruzzone, M. Buiatti, ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology* (2011) https:/doi.org/10.1111/j.1469-8986.2010.01061.x.

58. M. Rubinov, O. Sporns, Complex network measures of brain connectivity: Uses and interpretations. *Neuroimage* **52**, 1059–1069 (2010).

59. D. S. Bassett, *et al.*, Robust detection of dynamic community structure in networks. *Chaos* (2013) https:/doi.org/10.1063/1.4790830.