The relation between brain signal complexity and task difficulty on an executive function task

Article *in* NeuroImage · May 2019

CITATIONS
0

READS
8

4 authors, including:

John G Grundy
Iowa State University
25 PUBLICATIONS **176** CITATIONS
SEE PROFILE

John A. E. Anderson
Centre for Addiction and Mental Health
41 PUBLICATIONS **631** CITATIONS
SEE PROFILE

Judith M Shedden
McMaster University
46 PUBLICATIONS **623** CITATIONS
SEE PROFILE

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com

The relation between brain signal complexity and task difficulty on an executive function task

John G. Grundy [a, *], Ryan M. Barker [b], John A.E. Anderson [c], Judith M. Shedden [d]

[a] Department of Psychology, Iowa State University, Ames, IA, USA
[b] Department of Psychology, University of Toronto & Rotman Research Institute, Baycrest Hospital, Toronto, ON, Canada
[c] Krembil Family Imaging and Genetics Lab, Centre for Addiction and Mental Health, Toronto, ON, Canada
[d] Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, ON, Canada

ARTICLE INFO

Keywords:
Brain signal complexity
Multiscale entropy
Cognitive flexibility
Task-demands
Brain behavior relationships

ABSTRACT

On a daily basis, we constantly deal with changing environmental cues and perceptual conflicts and as such, our brains must flexibly adapt to current demands in order to act appropriately. Brains become more efficient and are able to switch states more readily by increasing the complexity of their neural networks. However, it is unclear how brain signal complexity relates to behavior in young adults performing cognitively demanding executive function tasks. Here we used multiscale entropy analysis and multivariate statistics on EEG data while participants performed a bivalency effect task-switching paradigm to show that brain signal complexity in young adults increases as task demands increase, that increases in brain signal complexity are associated with both speed gains and losses depending on scalp location, and that more difficult tasks are associated with more circumscribed complexity across the scalp. Overall, these findings highlight a critical role for brain signal complexity in predicting behavior on an executive function task among young adults.

1. Introduction

On a daily basis, we constantly deal with changing environmental cues and perceptual conflicts and as such, our brains must flexibly adapt to current demands. As brains develop, they increase both in efficiency and complexity (McIntosh et al., 2008; Mišic et al., 2010). These two concepts, while seemingly antithetical, actually pull in tandem. Brains *become* more efficient and are able to switch states more readily by increasing the complexity of their neural networks (Beharelle et al., 2012; Garrett et al., 2012; for reviews see Deco et al., 2011; Grady and Garrett, 2014).

Vitally, *complexity* should not be confused with random noise. By analogy, a piece of music can be complex, containing evolving melody, differing rhythms, and harmonies. Noise, by contrast, contains little to no structure and is dimensionally, and acoustically, less interesting. In a dynamic system, complexity can be useful in order to promote learning and exploration which in turn builds novel neural pathways. For example, Deco et al. (2011) suggest that brain signal complexity, or spontaneous fluctuations in signal, allows for training of neural networks. This spontaneous training strengthens the connections between neurons in these networks and leads to brain microstates. Because brain signal complexity continues to promote the spontaneous training of networks over time, these microstates will continue to become strengthened as they are re-explored. As a result, brain signal complexity allows individuals to switch between brain states more readily - greater complexity generates a greater number of network states that the brain can then visit. In line with this reasoning, Beharelle et al. (2012) showed that traumatic brain injury patients had less brain signal complexity than healthy controls, and that brain signal complexity was associated with greater accuracy and more stable responses on a visual feature-matching task. Importantly, the relationship between complexity and behavior was stronger for patients, suggesting that those with greater brain signal complexity recovered more of their cognitive ability after injury.

Sample entropy is one way to measure brain signal complexity (Richman and Moorman, 2000). Random noise has no structure or predictability - a single time point has no bearing on the following time point. In contrast, sample entropy does follow logic and the repetition of a signal can be modeled probabilistically. Sample entropy can be measured on neural time-series such as EEG or MEG to estimate the complexity (i.e. repeatability) of local and distributed networks mea-

* Corresponding author. Department of Psychology, Iowa State University, 901 Stange Rd., Ames, IA, 50011, USA.
Email address: grundy@iastate.edu (J.G. Grundy)

https://doi.org/10.1016/j.neuroimage.2019.05.045
Received 31 December 2018; Received in revised form 15 May 2019; Accepted 17 May 2019
Available online xxx
1053-8119/

sured across fine and coarse-grained timescales, respectively. As entropy can be measured at multiple time-scales (fine to coarse gradations), it is distinguished from sample entropy as "Multi-Scale-Entropy", or MSE (Costa et al., 2002, 2005). We chose to use MSE in the present study because it is more sensitive than other measures of complexity. Several studies, particularly in fMRI research, have used the standard deviation of the neural time series to estimate complexity, and this measure of complexity is often very informative for developmental outcomes (review in Garrett et al., 2013). However, consider the example offered by Pincus et al. (1991) regarding heart-rate variability, which illustrates why this metric may not always be ideal. Person 1's heart rate alternates predictably as 90, 70, 90, 70, 90, 70, 90, 70, 90, 70, whereas Person 2's heart rate fluctuates non-predictably as 90, 90, 70, 90, 70, 70, 70, 90, 90, 70. Both Person 1 and Person 2 have a mean of 80 and a standard deviation of 10 beats per minute. But the time series are not the same; Person 1's heart rate is predictable, whereas Person 2's heart rate is not, and it is known that sequences such as that of Person 2's heart rate are indicative of a healthier heart (Pincus et al., 1991). Thus, standard deviation is one measure of complexity (entropy), but it does not capture the full complexity of the signal.

Sample entropy, the measure of complexity used to create the multiple time-scales in MSE analyses, is a probabilistic metric of complexity that goes beyond standard deviations and can be distinguished from random noise. Put simply (see methods for expansion), the procedure creates a ratio using repetitions of N consecutive data points within a particular amplitude range compared to N+1 consecutive data points across the entire time series. More predictable sequences that produce smaller sample entropy values are deemed less complex. As mentioned previously, MSE takes this further by estimating the complexity across multiple time scales and can speak to local vs. distributed processing (Wang et al., 2018; Vakorin et al., 2011). For example, Wang et al. (2018) used both EEG and fMRI data to examine the relationship between functional connectivity and MSE and found that regional complexity was associated with greater functional connectivity with other distributed resting-state networks. We note that spectral power can also be used as a measure of complexity, but previous work has shown that it is not sensitive to certain dependencies within the neural time series to which MSE is sensitive (McIntosh et al., 2008). Furthermore, Sheehan et al. (2018) showed that sample entropy is independent of both spectral power and spectral slope in its predictability of memory performance.

MSE increases in a linear fashion across development at all timescales from ages 8 to 25 (McIntosh et al., 2008; Mišic et al., 2010), and this increase in complexity is associated with greater accuracy and less variable response times. As people enter old age, neural complexity patterns measured by MSE also change. Compared to young adults, elderly individuals rely more on local neural networks at fine scales, and less on distributed neural networks at coarse-grained scales (McIntosh et al., 2014; Heisz et al., 2015). There is some evidence that these shifts may be adaptive. When older adults shifted to local modes of processing they experienced performance gains. Evidence from young adults, however, is less clear. The shift in neural complexity experienced by older adults, where high performing individuals rely more on local networks than low performing individuals, was absent in a sample of younger adults performing a simple perceptual matching task (McIntosh et al., 2014). Thus, it is unclear how fine and coarse-grained brain signal complexity relates to behavior in young adults who are at peak cognitive performance.

MSE studies with young adults suggest that increasing neural complexity is associated with more accurate and less variable responses (McIntosh et al., 2008), but whether this relationship is driven by coarse or fine time-scales, or is undifferentiated, has not been explored since previous studies collapsed across time-scales. Furthermore, it is unclear whether the shift in complexity to more local neural networks, as is observed in older adults, would appear in younger adults in more demanding conditions. Past studies examining the relationship between signal complexity and behavior have relied on tasks such as the 1-back task or a very simple perceptual matching task (McIntosh et al., 2008; Mišic et al., 2010; McIntosh et al., 2014) that may not be cognitively demanding enough for young adults. As cognitive demands increase, younger adults may, like older adults, shift from more distributed, global processing, to rely more on local networks.

In the present study, young adult participants completed an executive function task with three trial types of varying difficulty while EEG was recorded in order to clarify the relationship between brain signal complexity, task difficulty, and behavior.

2. Methods

2.1. Participants

The data presented here were re-analyzed from those presented in Grundy et al. (2013). For clarity, with permission of the editors, we largely reproduce the description of the methods here, noting any additions or changes made for the present paper. Data from twenty-four young adult participants (mean age 20.30, SD = 3.48, seven males, 21 right-handed) from an event-related potential study were re-analyzed to examine brain signal complexity. All participants were first-year students recruited from McMaster University's Introductory Psychology undergraduate program and received course credit. Data from one participant was excluded from the original study due to an unknown corruption of the original file during PLS analysis (see below). This study was approved by the McMaster Research and Ethics Board and complied with the Tri-Agency Research Integrity Policy.

2.2. Task and procedure

The bivalency effect task switching paradigm (Grundy et al., 2013; Woodward et al., 2003; Meier et al., 2009) was used. This paradigm includes three trial types that vary in their level of difficulty: 1) *purely univalent* trials that cue only one task and appear in blocks with only univalent trials, 2) *univalent conflict* trials that are univalent trials appearing in a block where occasional conflicting bivalent trials appear, and 3) *bivalent* trials that cue two of the ongoing tasks even though only one of the features is relevant for performance (see Fig. 1). Several studies have replicated the robust finding that bivalent trials are the most difficult, followed by univalent conflict trials, then purely univalent trials (Grundy et al., 2013; Grundy and Shedden, 2014a, b; Meier and Rey-Mermet, 2012; Woodward et al., 2003, 2008). For each trial type, the primary outcome measures of interest were reaction times (ms) and accuracy (proportion correct).

Participants completed two practice blocks of 168 univalent trials in which individuals switched between three tasks (shape judgments: whether the shape was red or blue, case judgments: whether the letter was uppercase or lowercase, and parity judgments: whether the number was odd or even). This was followed by an identical pure block in which only univalent stimuli appeared, then by a block in which mostly univalent trials (152 of the 168) appeared with occasional bivalent stimuli (16 of the 168),[1] and finally by another purely univalent block to account for potential practice effects. The two pure univalent blocks were averaged together. The reason for the extensive practice during this paradigm is to ensure that participants were performing at ceiling on the easy trials before encountering the difficult bivalent tri-

[1] Given the limited number of bivalent trials, it was possible that outliers were contributing significantly to our findings. To address this issue, we identified any bivalent trials that fell above or below 2.5 standard deviations from the mean response time. Only one outlier was found (2.5 standard deviations below the mean). Importantly, exclusion of this participant did not significantly affect the results, and we therefore included all participants for completeness.

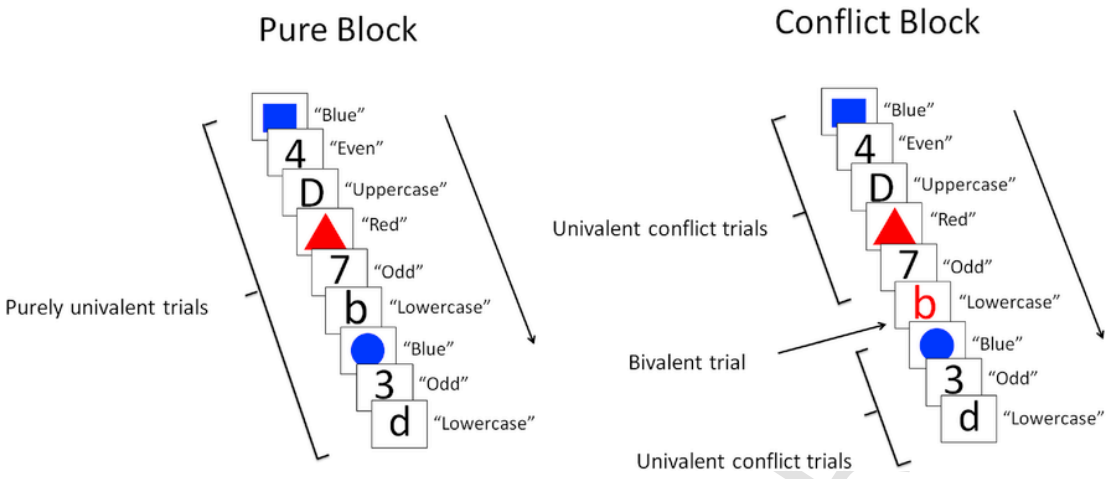

Fig. 1. Executive function bivalency effect paradigm used in the present study.

als. For other task parameters that are not the focus of this study, please see Grundy et al. (2013).

2.3. EEG collection

Scalp EEG data were collected with a 128-channel Biosemi ActiveTwo system (www.biosemi.com), with four additional electrodes placed at the outer canthi and just below each eye for recording horizontal and vertical eye movements. The sampling rate was set to 512 Hz, and the signals were bandpass filtered offline from 0.3 to 30 Hz. Signals were re-referenced to the common average reference. The baseline used was set to −200 ms prior to stimulus onset. Epochs of interest were between 0 and 900 ms after stimulus onset. Eye and motion artifacts were automatically identified and manually verified using the EEProbe software package (www.ant-neuro.com). Prototypes of eye movements were estimated for each individual and movement artifacts were subtracted across the electrode array. This was based on calculated vertical electro-oculogram (VEOG) propagation factors via a regression algorithm. Any trials with motion or eye movements that could not be corrected were removed from the analyses.

2.3.1. Brain signal complexity: multiscale entropy (MSE) analysis

We used MSE to assess brain signal complexity across multiple timescales. Fig. 2 illustrates how to calculate MSE. The following text is reprinted from the figure caption of Grundy et al. (2017), with permission from Elsevier: *A criterion length m (2 in the present study) defines a*

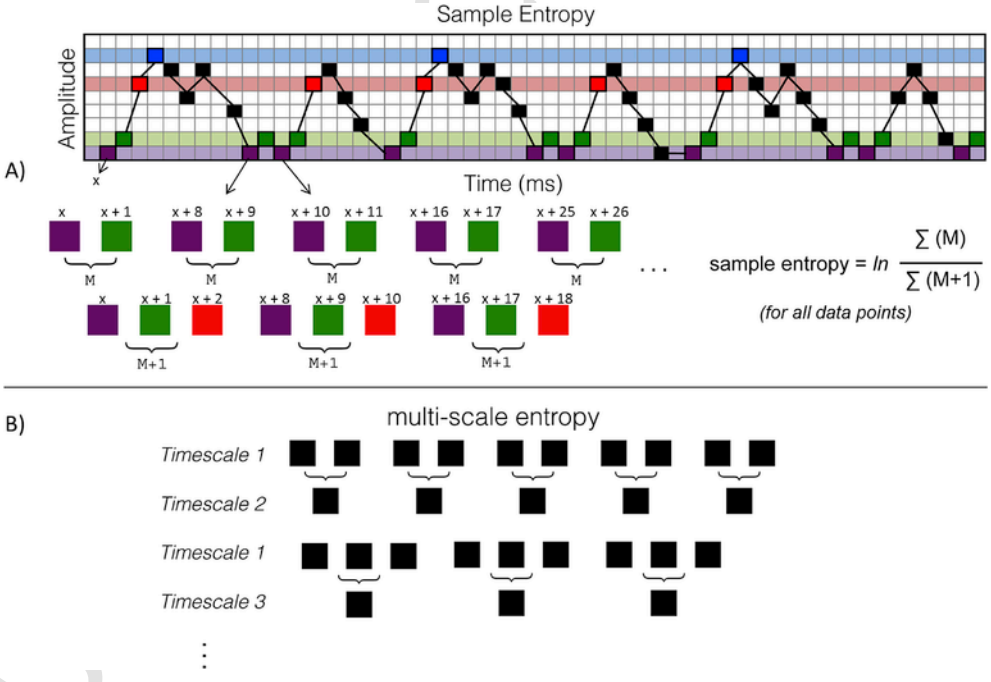

Fig. 2. *Reprinted from* Grundy et al. (2017). *Bilinguals have more complex EEG brain signals in occipital regions than monolinguals. NeuroImage, 159, 280–288, with permission from Elsevier.* Calculation of sample entropy and multiscale entropy using EEG brain signal. Panel A shows how to calculate sample entropy at each timescale. A criterion length m (2 in the present study) defines a number of consecutive data points in the EEG signal. For length m = 2 (starting with x and x + 1), sum the number of times that two consecutive data points occur within a pre-specified amplitude range (r = 0.5 in the present study) across time in the EEG waveform. Next, sum the number of times that three (m + 1) consecutive data points (starting with x, x + 1, and x + 2) occur within the amplitude range across time in the EEG waveform. This procedure is repeated for each data point in the time series (i.e., starting at x + 1, then starting at x + 2, then starting at x + 3, etc.) and all two-component matches and all three-component matches in the entire data series are summed together. Sample entropy is the natural logarithm of the ratio of two-component to three-component matches. Panel B represents downsampling of the original EEG data so that multiple fine to coarse timescales are created. This is done by averaging consecutive data points to create a new data point that corresponds to a coarser scale. This is done n times (in our study n = 20) and sample entropy is re-calculated at each timescale. This creates a series of fine to coarse grained entropy scales. This figure was inspired by Fig. 1 from Heisz et al. (2012).

number of consecutive data points in the EEG signal. For length m = 2 (starting with x and x + 1), sum the number of times that two consecutive data points occur within a pre-specified amplitude range (r = 0.5 in the present study) across time in the EEG waveform. Next, sum the number of times that three (m + 1) consecutive data points (starting with x, x + 1, and x + 2) occur within the amplitude range across time in the EEG waveform. This procedure is repeated for each data point in the time series (i.e., starting at x + 1, then starting at x + 2, then starting at x + 3, etc.) and all two-component matches and all three-component matches in the entire data series are summed together. Sample entropy is the natural logarithm of the ratio of two-component to three-component matches. Panel B represents downsampling of the original EEG data so that multiple fine to coarse timescales are created. This is done by averaging consecutive data points to create a new data point that corresponds to a coarser scale. This is done n times (in our study n = 20) and sample entropy is re-calculated at each timescale. This creates a series of fine to coarse grained entropy scales, which represent local to more distributed neural networks, respectively (Vakorin et al., 2011). Finally, in order to avoid 20 levels of the scale factor in our ANOVAs, we binned scales 1–6 as "fine", scales 7–14 as "mid", and scales 15–20 as "coarse".

We chose our MSE parameters based on our previous work (Grundy et al., 2017) and the work of others (Heisz et al., 2012) which showed these parameters to be sensitive to changes in brain signal complexity with task demands in young adults. A more detailed discussion of how to choose MSE parameters can be found in Lake et al. (2002). See Heisz and McIntosh (2013) for a nice video tutorial on how to apply MSE to EEG data.

For both behavioral and MSE data, an average of all trials for each participant for every condition was extracted for analysis.

2.4. Behavioral partial least squares analysis

Behavioral partial least squares analysis (PLS; see Krishnan et al., 2011) was used to derive multivariate, data-driven patterns of whole scalp entropy as it relates to behavioral measures. This technique decomposes the data into a set of orthogonal latent variables (LVs) in a single analytic step that maximally explains the covariance, in a manner similar to principal components analysis. Consequently, PLS is particularly well suited to the analysis of large data sets, such as neuroimaging data, that would otherwise suffer from loss of power due to corrections for multiple comparisons. We ran this analysis using the command line extension of the PLSgui software (Shen, 2009; McIntosh et al., 1996; McIntosh et al., 2004).

Conceptually, PLS is similar to canonical correlation, or principal components analysis (PCA), with a few exceptions. Most notably, PLS is particularly well suited to dealing with "wide" or rank deficient data (i.e., cases where there may be more variables, in this case electrodes, than participants). Another notable feature of PLS is that, while PCA will simply reduce the number of dimensions in the dataset, PLS attempts to find the linear combination of features that maximally covaries with the outcome measure (group, behavior etc.) Thus PLS is a "supervised" technique such that the algorithm attempts to fit a target that the user provides.

The inputs for behavioral PLS are brain and behavior matrices made up of stacked condition-wise submatrices. The brain matrix contains entropy observations (i.e., entropy values for each scale and electrode) across the columns, while rows represent participants within condition. The behavior matrix is similarly organized, and stores behavioral measures along the columns and participants within condition along the rows. Both matrices were normalized within condition and then combined through the multiplication of the transpose of the behavioral matrix and brain matrix for each condition submatrix. The result is a correlation matrix of stacked condition-wise submatrices. The rows of this matrix correspond to each behavior input and the columns depict the correlation between a given behavior and entropy observation from the original brain matrix.

Singular value decomposition is then performed on the resulting matrix of correlations, yielding orthogonal LVs that are organized in three outputs: brain saliences, design saliences, and a diagonal matrix of singular values. The brain salience represents patterns that are most strongly associated with the correlations found in the previous matrix. The original brain data is projected on to these values and summed to produce scalp scores. Scalp scores are values that indicate how strongly each participant is associated with the particular LV, and these may be averaged by condition to assess differences within the task.

Bootstrapping is performed to test the *reliability* of brain saliences. The original saliences are divided by the standard error calculated from the bootstrap sample to generate bootstrap ratios (BSRs), a standardized score the is analogous to a z-score (Efron and Tibshirani, 1986). Therefore, these BSRs may be considered reliable at the 0.05 level when they reach a value of two. This bootstrapping procedure also constructs bootstrapped confidence intervals around condition-wise scalp score averages which allows for the reliability of condition differences to be considered.

Permutation testing is also performed to assess the reliability of the overall LVs. This generates a distribution that is true under the null hypothesis. The original singular values are compared to these permuted distributions and the LV is considered significant if less than 5% of the new values exceed the original. For the current analysis, 500 bootstrap and permutation samples were used.

2.5. Exploratory correlation analyses

Finally, we conducted exploratory analyses to better understand how unified fluctuations in entropy were associated with each condition (pure univalent, univalent conflict, and bivalent) at every timescale (fine, mid, and coarse). We first computed pairwise Pearson correlations between all 128 electrodes across the scalp. To correct for multiple correlation tests, we corrected each of the matrices using the Benjamini-Hochberg method (8128 comparisons, q = .05; Benjamini and Hochberg, 1995). We next compared the matrices for each of the conditions within each scale bin using the Steiger test (Steiger, 1980). This method computes a sum of the squared difference between the two matrices. Similar matrices will necessarily produce values near zero while those that differ will generate a larger sum. This naturally forms a Chi-square distribution under the null hypothesis, and therefore a test statistic may be derived. To explore where the matrices differed for the most challenging condition, we compared corresponding correlations between univalent conflict and bivalent matrices using a dependent correlations test (Steiger's case B; Steiger, 1980). This accounts for the dependence of each variable in the two correlations being compared such that the covariance between all pairwise combinations is accounted for. We chose this method given the repeated design of the task and to also account for spatial correlation amongst electrodes. This resulted in three heatmaps that depicted significant correlation differences between univalent conflict and bivalent conditions for each scale bin.

3. Results

For transparency and reproducibility, all behavioral and MSE data used for the following analyses are uploaded to Figshare.com via the following link: https://figshare.com/articles/The_relation_between_brain_signal_complexity_and_task_difficulty_on_an_executive_function_task_data/8059907.

Unless otherwise stated, all post-hoc tests are Bonferroni corrected for multiple comparisons.

3.1. Behavioral performance

We first analyzed how task performance, response times and accuracy, varied as a function of condition using two one-way ANOVAs

(Fig. 3). For accuracy, a main effect of condition was observed, $F(2, 46) = 26.28$, $p < .001$, $\eta_p^2 = .53$. We found this effect was driven primarily by a relative drop in accuracy on bivalent trials, (i.e., accuracy was higher on pure univalent, $t(23) = 5.51$, $p < .001$, $d = 1.12$, and univalent conflict trials, $t(23) = 4.83$, $p < .001$, $d = 0.99$). Pure univalent and univalent conflict trials did not differ, $t(23) = 2.16$, $p = .12$, $d = 0.44$.

Similarly, response times revealed a main effect of condition, $F(2, 46) = 62.38$, $p < .001$, $\eta_p^2 = .73$. Post-hoc comparisons revealed that *all* conditions differed; participants were faster to respond to pure univalent trials than univalent conflict trials, $t(23) = -5.88$, $p < .001$, $d = -1.20$, and they responded to both univalent conflict trials, $t(23) = -7.62$, $p < .001$, $d = -1.56$, and pure univalent trials, $t(23) = -8.22$, $p < .001$, $d = -1.68$, faster than bivalent trials.

3.2. Multiscale entropy analyses

We investigated how brain signal complexity varied by condition (pure univalent, univalent conflict, bivalent) and entropy scale (fine, mid, coarse) using a two-way ANOVA (Fig. 4). The model revealed both main effects of condition, $F(2, 46) = 346.80$, $p < .001$, $\eta_p^2 = .93$, and scale, $F(2, 46) = 274.00$, $p < .001$, $\eta_p^2 = .92$.

Post-hoc comparisons revealed that entropy differed across all conditions, with higher entropy on univalent *conflict* than *pure* univalent trials, $t(23) = -7.92$, $p < .001$, $d = -1.62$, as well as higher entropy on *bivalent* trials than both univalent conflict, $t(23) = -16.17$, $p < .001$, $d = -3.30$, and pure univalent, $t = -23.96$, $p < .001$, $d = -4.89$, trials. We also found that entropy varied across all scale bins, with higher entropy on mid than fine scales, $t(23) = -17.73$, $p < .001$, $d = -3.62$, as well as higher entropy on coarse scales than both mid scales, $t(23) = -10.18$, $p < .001$, $d = -2.08$, and fine scales, $t(23) = -16.54$, $p < .001$, $d = -3.38$.

The main effects were modified by a significant interaction between condition and scale, $F(4, 92) = 168.10$, $p < .001$, $\eta_p^2 = .88$. Visual examination of Fig. 4 reveals that entropy is most similar across conditions at the fine scale, but diverges at the mid and coarse ranges. We then tested this observation empirically using two follow-up one-way ANOVAs which we ran on difference scores between adjacent scale bins. This has the effect of yielding a pseudo slope of entropy for each condition allowing us to compare the rate of entropy change between conditions as the scale increases.

We first compared the coarse and mid scales by condition (see Fig. 4b), which approached, but did not meet conventional levels of statistical significance, $F(2, 46) = 3.83$, $p = .058$, $\eta_p^2 = .15$, suggesting a greater difference between these scales in the bivalent than the pure univalent condition, $t(23) = -2.46$, $p = .07$, $d = -0.50$, but not between the bivalent and univalent conflict or between the univalent conflict and pure univalent conditions, all ts < 2.0. We describe the pattern here to highlight its similarity to the next analysis, but note that these statistics did not reach conventional levels of significance and caution any over-interpretation.

Our second model examined how entropy changed across the mid and fine scales by condition (see Fig. 4c). The main effect of condition was significant, $F(2, 46) = 313.70$, $p = < .001$, $\eta_p^2 = .93$. Post-hoc tests revealed that the difference between mid and fine-scale entropy differed across *all* conditions. In increasing order of magnitude, the univalent conflict trials had a greater difference between entropy scales than the pure univalent condition, $t(23) = -6.49$, $p < .001$, $d = -1.33$, and the bivalent condition had a greater change in entropy than the univalent conflict trials, $t(23) = -16.65$, $p = < .001$, $d = -3.40$.

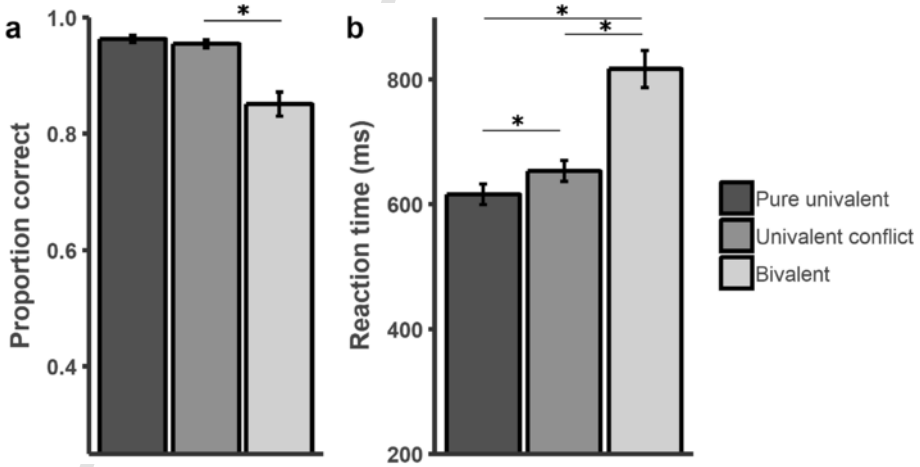

Fig. 3. Accuracy (a) and reaction times (b) plotted by condition. Error bars represent standard error.

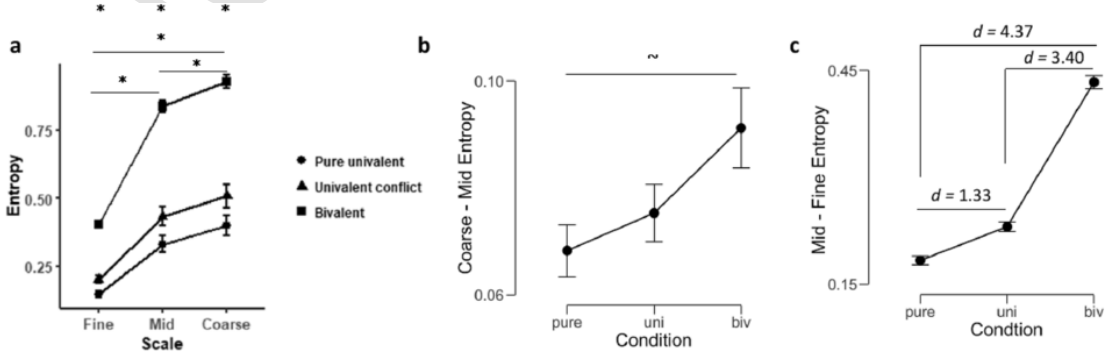

Fig. 4. Level of entropy by scale and condition. Error bars represent standard errors. *$p < 0.05$, ~ $p = 0.058$.

This pattern of findings suggested that task difficulty was associated with increased entropy from fine to mid scales (see Fig. 4c), but that the most effortful condition, bivalent, was associated with a much larger effect (see inset Cohen's *d* measures for relative difference between conditions).

3.3. Individual performance and entropy

To examine whether greater levels of complexity at fine scales are related to better performance in young adults when task demands are sufficiently high, we ran a three-way ANOVA in which we compared condition, scale, and individual performance (high vs low; calculated using a median split of efficiency scores (RT/proportion correct) that account for speed-accuracy tradeoffs (Townsend and Ashby, 1983).

Replicating our earlier findings, there was a main effect of condition, $F(2, 44) = 348.04$, $p = < .001$, $\eta_p^2 = .94$, and all conditions differed, (all *t*s > 7.50). We also found an effect of scale, $F(2, 44) = 269.18$, $p = < .001$, $\eta_p^2 = .92$, with entropy greatest at coarse scales and lowest at fine scales (all *t*s > 10.00). Moreover, the interaction between condition and scale remained significant, $F(4, 88 = 164.70$, $p = < .001$, $\eta_p^2 = .88$. Critically, the effect of individual performance was *not* significant, $F(1, 22) = 0.13$, $p = .72$, $\eta_p^2 = .01$, nor were any interactions with this variable (all *F*s < 1.10), suggesting that individual performance was not associated with the trajectory of multi-scale entropy on this cognitively demanding task.

3.4. Partial least squares analysis

We next ran a multivariate behavioral PLS analysis in order to assess how entropy, condition (e.g., univalent, univalent conflict, bivalent), and behavior covaried across all scales. As noted earlier, PLS computes these relationships in a single step, thus obviating the need for multiple comparison correction. Reliable covariance patterns between condition, entropy, and behavior were assessed with bootstrap ratios (BSRs, see Fig. 5a). We also projected BSRs onto topographic electrode space for fine, medium, and coarse scales to aid interpretation. Negative BSRs represent regions where entropy is associated with slower response times and positive BSRs are associated with faster response times across all three conditions.

The analysis reduced the dimensionality of the dataset to a single significant LV that maximally explained brain-behavior relationships. Negative BSRs formed the more prominent pattern, comprising most of the scalp topography, with both left and right frontotemporal as well as left and midline posterior sites most represented, across all time scales. Positive BSRs were found to occupy left-lateralized frontal electrodes during fine and middle time scales. Scalp score loadings by condition and behavior (Fig. 5b) indicated that accuracy and RT loaded opposite to each other such that there was tendency for faster response times and more accurate responses in left frontal regions (warm regions) for fine and mid-scale entropy, but slower and less accurate responses across various locations across the scalp at all timescales (cool regions). However, accuracy was not reliably associated with entropy produced

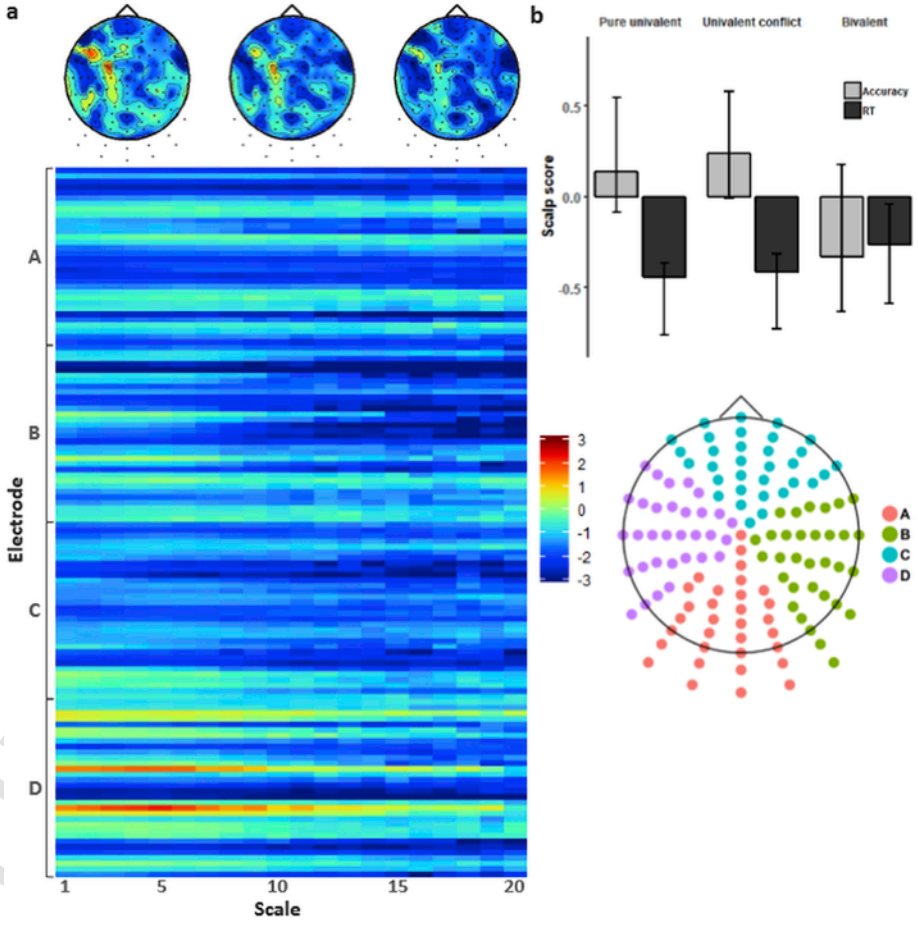

Fig. 5. Bootstrap ratios across electrode and time scale (a). Scalp score loadings across all electrodes. The topographies at the top represent the binned (fine, mid, and coarse) scales across the scalp. (b) indicate how BSR patterns correlate with behavioral performance by condition. Negative BSRs represent regions where entropy is associated with slower response times and positive BSRs are associated with faster response times across all three conditions. Error bars represent bootstrapped 95% confidence intervals.

by the analysis given that the confidence intervals overlapped with zero.

3.5. Exploratory analysis: task-based entropy coupling

While Bonferroni adjustments were applied to correct for multiple comparisons in the other manuscript sections, the False-Discovery-Rate correction methods was used here due to the large number of correlations and the exploratory nature of the analyses in this section.

We investigated entropy region-to-region coupling across condition and scale bin using thresholded correlation matrices (see Figs. 6–8). Across these matrices, the predominant pattern that emerged was that less effortful *univalent* conditions were associated with widespread positive coupling between electrodes compared to the more challenging bivalent condition which was associated with a much sparser pattern of covariance. Univalent conditions yielded a dense pattern of reliable correlations across the scalp. By contrast, the bivalent condition was associated with only a few reliable clusters of correlations. This pattern was replicated across all three scale bins. The major clusters of connectivity within the bivalent condition that were revealed from the analysis were between occipital and right parieto occipito electrodes, parietal midline, and right-mid parietal electrodes, as well as correlations within right temporal lateral electrodes, and finally within occipital electrodes.

Differences between the conditions were tested by comparing each correlation matrix within each scale bin (see Table 1).

This pattern of results confirms that while both *univalent* conditions were associated with similarly strong patterns of whole-scalp coupling, the more effortful bivalent condition drastically altered this pattern, reducing widespread connectivity to more constrained regional coupling. To provide a spatial representation of where differences were most prominent, the difference between the univalent conflict and bivalent conditions for each pairwise correlation was tested. The result is a heatmap depicting where reliable differences are found between the original condition matrices (Fig. 9). Differences were most reliable between these conditions at mid and coarse scales in correlations between right temporal and parieto occipital sites.

4. Discussion

In the present study, we sought to clarify the relationship between brain signal complexity and behavior in young adult participants performing an executive function bivalency effect task (Grundy et al., 2013). Several important findings emerged. 1) Brain signal complexity was greatest in the most difficult conditions. 2) Across all trial types that varied in level of difficulty and across all timescales, greater brain signal complexity predicted faster and slower response times depending on scalp location, but signal complexity was not associated with accuracy. 3) Contrary to our hypothesis, better performing individuals did *not* show a shift from coarse to fine scale entropy on more difficult tasks. As task difficulty increased, however, brain signal complexity instead *decoupled* across the scalp. These findings have important implications for how brain signal complexity relates to behavior in young adults.

Brain signal complexity increased in a linear fashion as task difficulty increased, with greatest signal complexity for the most difficult bivalent trials and least signal complexity for the easiest purely univalent trials. To our knowledge, this is the first time that multiscale entropy has been related to task difficulty on an executive function task. The pattern was most reliable for fine and mid timescales, reflecting more local than distributed neural network processing. These findings might be explained by previous work showing that brain signal complexity might reflect greater knowledge representations. For example, Heisz et al. (2012) found that brain signal complexity was greater for familiar than unfamiliar faces and that learning faces also led to greater brain signal complexity. Similarly, Mišic' et al. (2010) found that upright faces were more complex than inverted faces and upright faces presumably have relational properties between features that inverted faces do not. Bivalent trials in the present study contained two types of information, task-relevant features, and task-irrelevant features

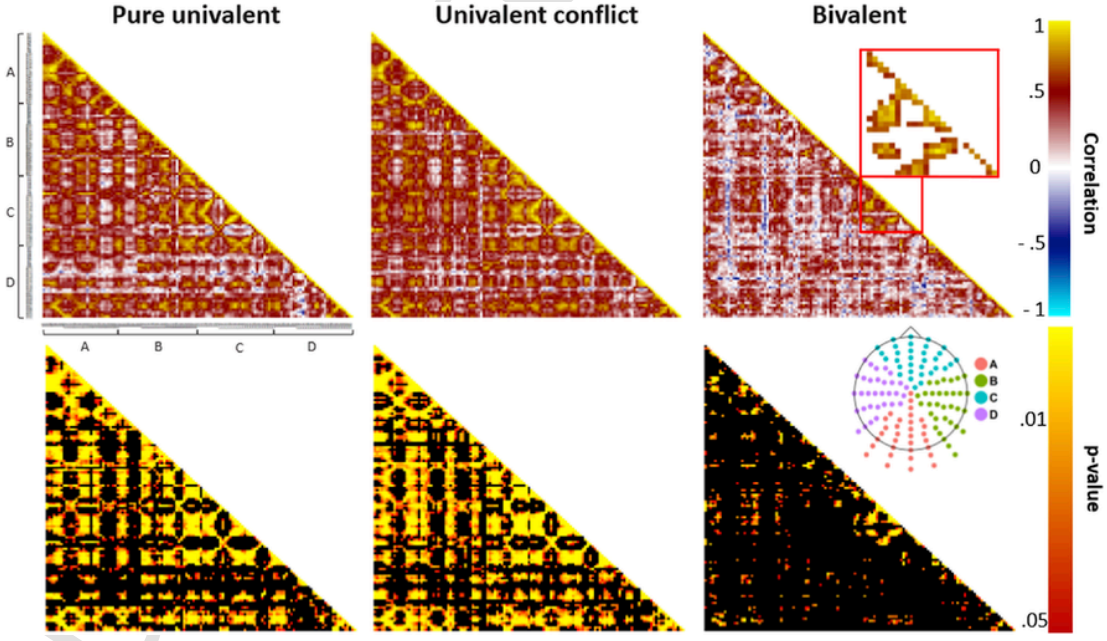

Fig. 6. Correlation between electrodes at fine scales. Raw *r* values appear in the top row while corrected p-values are reported in the bottom row. A, B, C, and D refer generally to posterior, right, anterior, and left electrode scalp sites. Colors in the top panel represent positive (warm colors) and negative (cool colors) correlations. White represents no correlation. The largest reliable cluster grouping across timescales for the bivalent condition is highlighted in the magnified section. Colors in the bottom panel represent the significance of the *p*-value associated with the correlations, from 0.05 (red) to < 0.01 (yellow). Black represents non-significant *p*-values > 0.05. The correlations in the top panel represent the full range of uncorrected correlations whereas the *p*-values associated with those correlations in the bottom panel are thresholded at an FDR $p < 0.05$ level. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

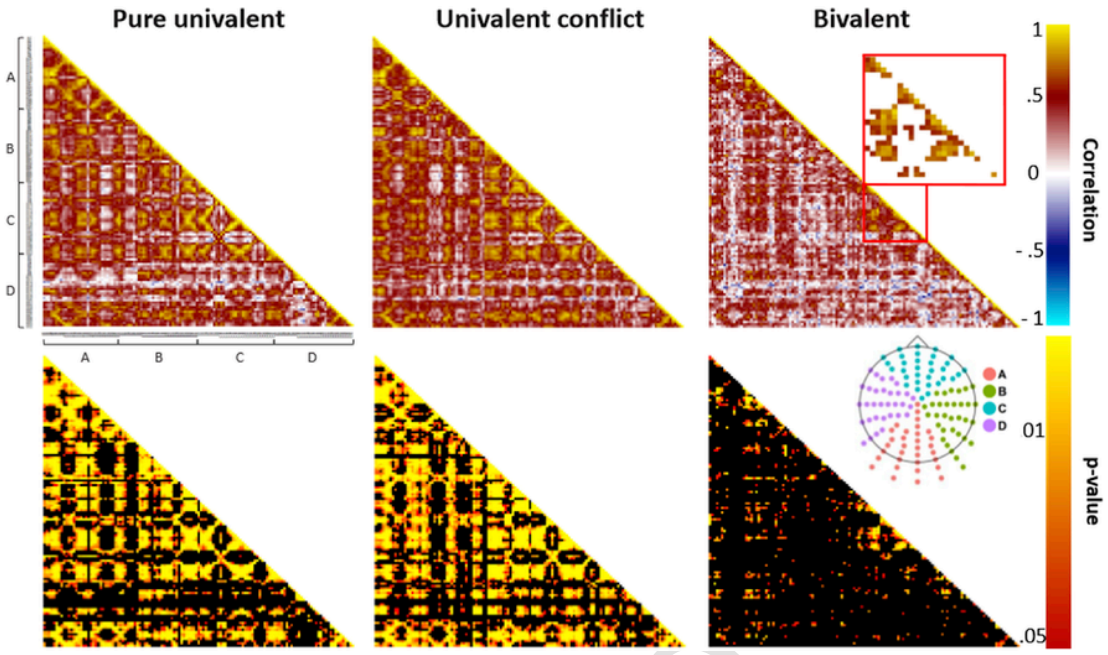

Fig. 7. Correlation between electrodes at mid scales. Raw *r* values appear in the top row while corrected p-values are reported in the bottom row. A, B, C, and D refer generally to posterior, right, anterior, and left electrode scalp sites. Colors in the top panel represent positive (warm colors) and negative (cool colors) correlations. White represents no correlation. The largest reliable cluster grouping across timescales for the bivalent condition is highlighted in the magnified section. Colors in the bottom panel represent the significance of the *p*-value associated with the correlations, from 0.05 (red) to < 0.01 (yellow). Black represents non-significant *p*-values < 0.05. The correlations in the top panel represent the full range of uncorrected correlations whereas the *p*-values associated with those correlations in the bottom panel are thresholded at an FDR $p < 0.05$ level. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

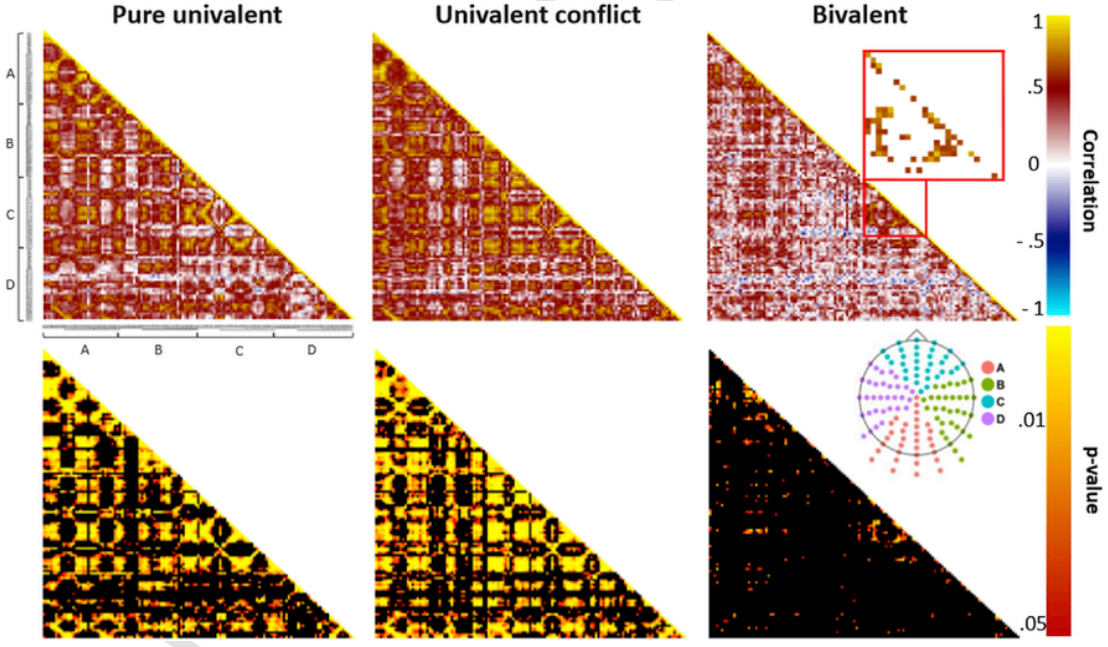

Fig. 8. Correlation between electrodes at coarse scales. Raw *r* values appear in the top row while corrected p-values are reported in the bottom row. A, B, C, and D refer generally to posterior, right, anterior, and left electrode scalp sites. Colors in the top panel represent positive (warm colors) and negative (cool colors) correlations. White represents no correlation. The largest reliable cluster grouping across timescales for the bivalent condition is highlighted in the magnified section. Colors in the bottom panel represent the significance of the *p*-value associated with the correlations, from 0.05 (red) to < 0.01 (yellow). Black represents non-significant *p*-values < 0.05. The correlations in the top panel represent the full range of uncorrected correlations whereas the *p*-values associated with those correlations in the bottom panel are thresholded at an FDR $p < 0.05$ level. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

and thus had greater knowledge representations than purely univalent trials with only one task-relevant feature.

The finding that greater brain signal complexity was associated with faster performance in left-lateralized frontal regions of the scalp is roughly in line with previous research from McIntosh et al. (2008) who showed a strong positive relationship between signal complexity and performance. However, across multiple regions of the scalp (the negative BSR values in Fig. 5) and across all timescales, brain signal complexity was associated with *slower* response times as well. Furthermore, unlike McIntosh et al., brain signal complexity in the present study was

Table 1
Chi-Squared significance testing results for differences between correlation matrices within each scale bin in the exploratory analysis of task-based entropy coupling.

	Fine	Mid	Coarse
Bivalent vs.Pure Univalent	$\chi^2(8128) = 111358.86$, $p < .001$	$\chi^2(8128) = 10174.38$, $p < .001$	$\chi^2(8128) = 10400.02$, $p < .001$
Bivalent vs. Univalent Conflict	$\chi^2(8128) = 11156.31$, $p < .001$	$\chi^2(8128) = 10772.78$, $p < .001$	$\chi^2(8128) = 12771.12$, $p < .001$
Univalent Conflict vs. Pure Univalent	$\chi^2(8128) = 3175.10$, $p > .05$	$\chi^2(8128) = 3668.91$, $p > .05$	$\chi^2(8128) = 4290.20$, $p > .05$

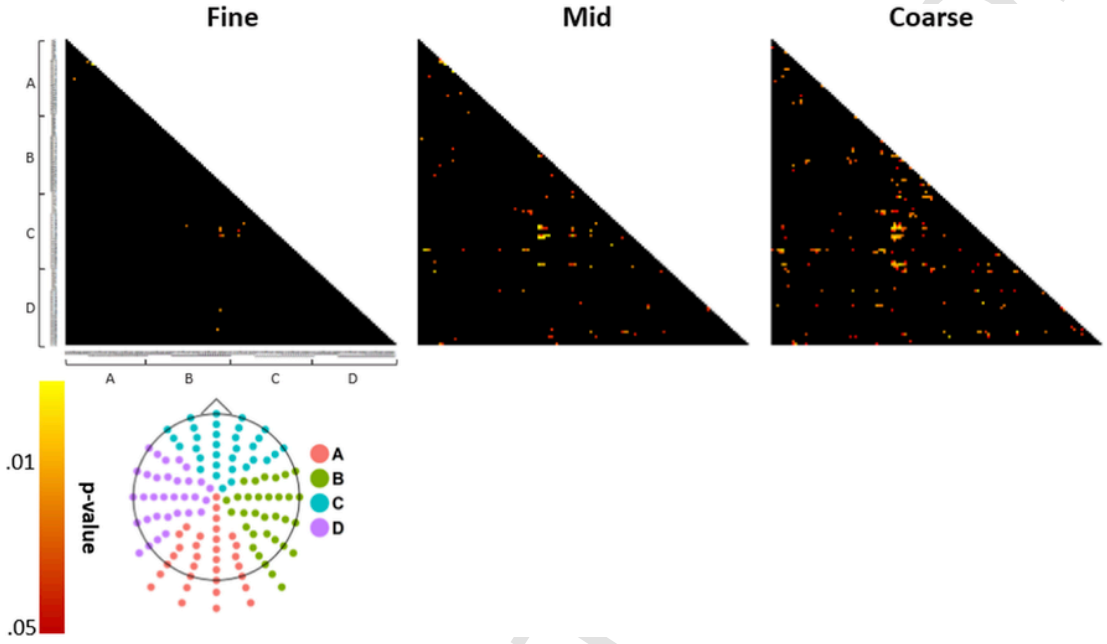

Fig. 9. Corrected correlation differences between univalent conflict and bivalent conditions at each scale bin. A, B, C, and D refer generally to posterior, right, anterior, and left electrode scalp sites. Colors represent the significance of the *p*-value associated with the correlations, from 0.05 (red) to < 0.01 (yellow). Black represents non-significant *p*-values < 0.05. Correlations are thresholded at an FDR $p < 0.05$ level. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

associated with response times rather than accuracy. This was surprising, but it is possible that the nature of the task being used dictated this shift. McIntosh et al. used a one-back working memory task that is not as cognitively demanding as the task-switching paradigm that we used. Consider that a one-back task requires only a decision about whether the last picture was the same as the current picture and requires very little demands on working memory, with no need to update working memory as you would in a more demanding two-back task. On the other hand, the paradigm used in our study required individuals to hold in mind task-sets for three separate tasks. Furthermore, bivalent trials require one to focus attention on task-relevant features while ignoring task-irrelevant features and this is cognitively demanding. The change in task difficulty may have led to a shift from brain signal complexity relating to accuracy to brain signal complexity relating to speed. It is also possible that this shift is a result of using a continuous memory updating task in one case and a continuous task-switching paradigm in the other, rather than solely cognitive demand. Thus, it may be a qualitative rather than a quantitative difference. While the shift from accuracy to RT still requires explanation and warrants further investigation, it highlights how brain signal complexity's relationship with behavior depends on the nature of the task.

Contrary to our hypothesis, high performers and low performers had similar brain signal complexity profiles. That is, there was no shift for high performers to show more fine scale and less coarse scale entropy than low performers. This finding replicates the results of McIntosh et al. (2014) who, again, used a relatively less cognitively demanding task. Thus, using more local neural networks does not appear to benefit young adult performance, even when task demands are increased. It is possible that a shift to more local neural networks in older adult populations may not be indicative of healthy aging, as previously suggested (McIntosh et al., 2014), but rather a requirement as the brain starts to deteriorate. Those who do not make this shift may perform worse behaviorally, explaining previous findings in which low performing elderly show less fine-scale and more coarse scale entropy than high performing elderly. It must be noted however that Heisz et al. (2015) showed that older adults who are more physically active also show a shift to more local complexity than less physically active individuals and physical activity is believed to slow cognitive decline (see Sofi et al., 2011 for a meta-analysis of prospective studies). It is possible that physical activity allows one to make the required shift from global to local information processing earlier, but such a claim requires further investigation.

Although it is clear that our low performers and high performers did not differ in terms of brain signal complexity profiles, our exploratory analysis examining the relationship between brain signal complexity across the scalp revealed a pattern that may be in line with more local processing as task demands are increased. On the easiest tasks, brain signal complexity was highly correlated in an undifferentiated manner across most electrodes. By contrast, on the most difficult task where conflict was present (the bivalent trials), brain signal complexity was far more circumscribed. In short, easier conditions were associated with greater overall connectivity, and less specificity between regions whereas conditions requiring focused control were associated with localized, modular, coupling between regions in the parietal and occipital scalp regions. In graph-theory, a modular brain organization can reflect specialized processing, and higher modularity tends to be a characteristic of younger, and healthier, functional brain networks (Arnemann et al., 2015; e.g., Brier et al., 2014). Relative to young adults, older adults tend to experience a breakdown of the functional units (modules), as functional regions bleed into one another and be-

come dedifferentiated (Zheng et al., 2018) while at the same time, large-scale functional networks also start to decouple (Andrews-Hanna et al., 2007). There is some suggestion that this dedifferentiation might be compensatory (Cabeza, 2002; Davis et al., 2008), however, other evidence suggests that older people who maintain a more "youthful" brain organization also perform better on cognitive tasks (Grady et al., 2010). Thus while our original hypothesis which predicted a global shift from coarse to fine entropy as task demand increased was not borne out in the data, our subsequent exploratory analyses of correlation organization seem to support the shift to local, highly connected modules during demanding tasks. Our two findings are not necessarily incompatible since "global" and "local" processing are defined differently. Our initial hypothesis focused on global versus local processing as intrinsic properties of coarse and fine scales. By contrast, our follow-up exploratory analysis tested global and local properties more directly by comparing the sparsity of the correlation matrices at varying scales. Our findings point to functional reorganization in young adults as task difficulty increases, though we note that this should be independently replicated given the exploratory nature of the analysis.

5. Conclusion

Brain signal complexity is increasingly becoming a useful metric by which to examine network efficiency (Beharelle et al., 2012; Deco et al., 2011), knowledge representations (Heisz et al., 2012), and distributed vs. local neural network processing (McIntosh et al., 2014; Vakorin et al., 2011). Here we add to the growing body of literature by showing that brain signal complexity in young adults increases as task demands increase, that increases in brain signal complexity are associated with both speed gains and losses depending on scalp location and that more difficult tasks are associated with more circumscribed complexity across the scalp. Overall, these findings highlight a critical role for brain signal complexity in predicting behavior on an executive function task among young adults.

Acknowledgments

The authors declare no conflicts of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2019.05.045.

References

Andrews-Hanna, J.R., Snyder, A.Z., Vincent, J.L., Lustig, C., Head, D., Raichle, M.E., Buckner, R.L., 2007. Disruption of large-scale brain systems in advanced aging. Neuron 56 (5), 924–935.

Arnemann, K.L., Chen, A.J.-W., Novakovic-Agopian, T., Gratton, C., Nomura, E.M., D'Esposito, M., 2015. Functional brain network modularity predicts response to cognitive training after brain injury. Neurology 84 (15), 1568–1574.

Beharelle, A.R., Kovačević, N., McIntosh, A.R., Levine, B., 2012. Brain signal variability relates to stability of behavior after recovery from diffuse brain injury. Neuroimage 60, 1528–1537.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B 57, 289–300.

Brier, M.R., Thomas, J.B., Fagan, A.M., Hassenstab, J., Holtzman, D.M., Benzinger, T.L., et al., 2014. Functional connectivity and graph theory in preclinical Alzheimer's disease. Neurobiol. Aging 35 (4), 757–768.

Cabeza, R., 2002. Hemispheric asymmetry reduction in older adults: the HAROLD model. Psychol. Aging 17 (1), 85–100.

Costa, M., Goldberger, A.L., Peng, C.K., 2002. Multiscale entropy analysis of complex physiologic time series. Phys. Rev. Lett. 89, 068102.

Costa, M., Goldberger, A.L., Peng, C.K., 2005. Multiscale entropy analysis of biological signals. Phys. Rev. 71, 021906.

Davis, S.W., Dennis, N.A., Daselaar, S.M., Fleck, M.S., Cabeza, R., 2008. Que PASA? The posterior–anterior shift in aging. Cerebr. Cortex 18 (5), 1201–1209.

Deco, G., Jirsa, V.K., McIntosh, A.R., 2011. Emerging concepts for the dynamical organization of resting-state activity in the brain. Nat. Rev. Neurosci. 12, 43–56.

Efron, B., Tibshirani, R., 1986. Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy. Stat. Sci. 1, 54–77.

Garrett, D.D., Kovacevic, N., McIntosh, A.R., Grady, C.L., 2012. The modulation of BOLD variability between cognitive states varies by age and processing speed. Cerebr. Cortex 23, 684–693.

Garrett, D.D., Samanez-Larkin, G.R., MacDonald, S.W., Lindenberger, U., McIntosh, A.R., Grady, C.L., 2013. Moment-to-moment brain signal variability: a next frontier in human brain mapping?. Neurosci. Biobehav. Rev. 37 (4), 610–624.

Grady, C.L., Garrett, D.D., 2014. Understanding variability in the BOLD signal and why it matters for aging. Brain Imag. Behav. 8, 274–283.

Grady, C.L., Protzner, A.B., Kovacevic, N., Strother, S.C., Afshin-Pour, B., Wojtowicz, M., et al., 2010. A multivariate analysis of age-related differences in default mode and task-positive networks across multiple cognitive domains. Cerebr. Cortex 20 (6), 1432–1447.

Grundy, J.G., Anderson, J.A.E., Bialystok, E., 2017. Bilinguals have more complex EEG brain signals in occipital regions than monolinguals. Neuroimage 159, 280–288.

Grundy, J.G., Benarroch, M.F., Woodward, T.S., Metzak, P.D., Whitman, J.C., Shedden, J.M., 2013. The Bivalency effect in task switching: event-related potentials. Hum. Brain Mapp. 34, 999–1012.

Grundy, J.G., Shedden, J.M., 2014a. A role for recency of response conflict in producing the bivalency effect. Psychol. Res. 78, 679–691.

Grundy, J.G., Shedden, J.M., 2014b. Support for a history-dependent predictive model of dACC activity in producing the bivalency effect: an event-related potential study. Neuropsychologia 57, 166–178.

Heisz, J.J., Gould, M., McIntosh, A.R., 2015. Age-related shift in neural complexity related to task performance and physical activity. J. Cogn. Neurosci. 27, 605–613.

Heisz, J.J., McIntosh, A.R., 2013. Applications of EEG neuroimaging data: event-related potentials, spectral power, and multiscale entropy. J. Vis. Exp. 76, 50131.

Heisz, J.J., Shedden, J.M., McIntosh, A.R., 2012. Relating brain signal variability to knowledge representation. Neuroimage 63, 1384–1392.

Krishnan, A., Williams, L.J., McIntosh, A.R., Abdi, H., 2011. Partial least squares (PLS) methods for neuroimaging: a tutorial and review. Neuroimage 46 (2), 455–475. https://doi.org/10.1016/j.neuroimage.2010.07.034.

Lake, D.E., Richman, J.S., Griffin, M.P., Moorman, J.R., 2002. Sample entropy analysis of neonatal heart rate variability. Am. J. Physiol. Regul. Integr. Comp. Physiol. 283, R789–R797.

McIntosh, A.R., Kovacevic, N., Itier, R.J., 2008. Increased brain signal variability accompanies lower behavioral variability in development. PLoS Comput. Biol. 4, e1000106.

McIntosh, A.R., Bookstein, F.L., Haxby, J.V., Grady, C.L., 1996. Spatial pattern analysis of functional brain images using partial least squares. Neuroimage 3, 143–157.

McIntosh, A.R., Chau, W.K., Protzner, A.B., 2004. Spatiotemporal analysis of event-related fMRI data using partial least squares. Neuroimage 23, 764–775.

McIntosh, A.R., Vakorin, V., Kovacevic, N., Wang, H., Diaconescu, A., Protzner, A.B., 2014. Spatiotemporal dependency of age-related changes in brain signal variability. Cerebr. Cortex 24, 1806–1817.

Meier, B., Rey-Mermet, A., 2012. Beyond feature binding: interference from episodic context binding creates the bivalency effect in task-switching. Front. Psychol. 3, 386.

Meier, B., Woodward, T.S., Rey-Mermet, A., Graf, P., 2009. The bivalency effect in task switching: general and enduring. Can. J. Exp. Psychol. 63, 201–210.

Mišic, B., Mills, T., Taylor, M.J., McIntosh, A.R., 2010. Brain noise is task dependent and region specific. J. Neurophysiol. 104, 2667–2676.

Pincus, S.M., Gladstone, I.M., Ehrenkranz, R.A., 1991. A regularity statistic for medical data analysis. J. Clin. Monit. 7, 335–345.

Richman, J.S., Moorman, J.R., 2000. Physiological time-series analysis using approximate entropy and sample entropy. Am. J. Physiol. Heart Circ. Physiol. 278, H2039–H2049.

Sheehan, T.C., Sreekumar, V., Inati, S.K., Zaghloul, K.A., 2018. Signal complexity of human intracranial EEG tracks successful associative-memory formation across individuals. J. Neurosci. 38, 1744–1755.

Shen, J., 2009. PLSgui User's Guide, Retrieved from http://www.rotman-baycrest.on.ca/pls/UserGuide.htm.

Sofi, F., Valecchi, D., Bacci, D., Abbate, R., Gensini, G.F., Casini, A., Macchi, C., 2011. Physical activity and risk of cognitive decline: a meta-analysis of prospective studies. J. Intern. Med. https://doi.org/10.1111/j.1365-2796.2010.02281.x.

Steiger, J.H., 1980. Tests for comparing elements of a correlation matrix. Psychol. Bull. 87 (2), 245–251.

Townsend, J.T., Ashby, F.G., 1983. Stochastic Modeling of Elementary Psychological Processes. Cambridge University Press, New York.

Vakorin, V.A., Lippe, S., McIntosh, A.R., 2011. Variability of brain signals processed locally transforms into higher connectivity with brain development. J. Neurosci. 31, 6405–6413.

Wang, D.J., Jann, K., Fan, C., Qiao, Y., Zang, Y.F., Lu, H., Yang, Y., 2018. Neurophysiological basis of multi-scale entropy of brain complexity and its relationship with functional connectivity. Front. Neurosci. 12, 352.

Woodward, T.S., Meier, B., Tipper, C., Graf, P., 2003. Bivalency is costly: bivalent stimuli elicit cautious responding. Exp. Psychol. 50, 233–238.

Woodward, T.S., Metzak, P.D., Meier, B., Holroyd, C.B., 2008. Anterior cingulate cortex signals the requirement to break inertia when switching tasks: a study of the bivalency effect. Neuroimage 40, 1311–1318.

Zheng, L., Gao, Z., Xiao, X., Ye, Z., Chen, C., Xue, G., 2018. Reduced fidelity of neural representation underlies episodic memory decline in normal aging. Cerebr. Cortex 28 (7), 2283–2296.