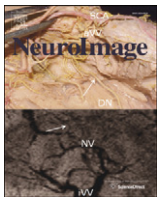
NeuroImage 63 (2012) 1384–1392

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg



Relating brain signal variability to knowledge representation

Jennifer J. Heisz [a,b,*], Judith M. Shedden [b], Anthony R. McIntosh [a,*]

[a] Rotman Research Institute at Baycrest Center, University of Toronto, Toronto, Canada
[b] Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, Canada

ARTICLE INFO

Article history:
Accepted 5 August 2012
Available online 11 August 2012

Keywords:
Multiscale entropy
Familiarity
Repetition effects
Complexity
EEG
Information theory

ABSTRACT

We assessed the hypothesis that brain signal variability is a reflection of functional network reconfiguration during memory processing. In the present experiments, we use multiscale entropy to capture the variability of human electroencephalogram (EEG) while manipulating the knowledge representation associated with faces stored in memory. Across two experiments, we observed increased variability as a function of greater knowledge representation. In Experiment 1, individuals with greater familiarity for a group of famous faces displayed more brain signal variability. In Experiment 2, brain signal variability increased with learning after multiple experimental exposures to previously unfamiliar faces. The results demonstrate that variability increases with face familiarity; cognitive processes during the perception of familiar stimuli may engage a broader network of regions, which manifests as higher complexity/variability in spatial and temporal domains. In addition, effects of repetition suppression on brain signal variability were observed, and the pattern of results is consistent with a selectivity model of neural adaptation.

Introduction

It is becoming clear that transient fluctuation in brain signal – i.e., variability – conveys important information about network dynamics that cannot be obtained from mean brain activity alone (Garrett et al., 2011; Vakorin et al., 2011). Within the context of simulated neural networks, information integration across widespread neural networks is achieved through the emergence and disappearance of correlated activity between network nodes over time and across multiple timescales (Deco et al., 2011; Tononi et al., 1998). Such transient changes in the global functional connectivity pattern cause fluctuations in the temporal dynamics of the corresponding brain signal (Breakspear, 2002; Freeman and Rogers, 2002; Friston, 2001; Honey et al., 2007); networks with more potential configurations have a greater repertoire to dynamically explore and elicit a more variable response (Deco et al., 2011; Ghosh et al., 2008; Tsuda, 2001). In turn, signal variability may reflect the information processing capacity of the system where higher variability would indicate greater information integration across networks via dynamical network reconfiguration. Indeed, previous empirical work suggests that signal variability increases with the amount of information available for a stimulus (Misic et al., 2010). It follows that variability should increase with learning as new information is acquired. Namely, a stimulus with greater acquired information content should produce a brain signal that carries more information.

To illustrate the effect that learning may have on brain signal variability, consider how familiarity with a face would affect the dynamics of the face-processing network. The face-processing network consists of a core visual system involved in facial image processing plus an extended system involved in processing of associated semantic information (e.g., name, occupation; Haxby et al., 2000). The presentation of a face activates the core visual system coding for the image itself. As a face becomes familiar, information builds up in both the core visual and extended regions (Burton et al., 1999), which increases the number of potential functional configurations the network can occupy. That is, familiarity increases the repertoire of responses produced by the brain when the face is presented. Consequently, highly familiar faces associated with greater information content activate networks with greater repertoire to produce a more variable response (Deco et al., 2011; Tononi et al., 1998).

The present study examined whether the variability of human EEG reflects the information available for face processing. We employed the information theoretic metric (Gatlin, 1972; Shannon, 1948) multiscale entropy (MSE; Costa et al., 2005, Heisz & McIntosh, in press), which uses sample entropy (Richman and Moorman, 2000) to estimate the variability of the neuroelectical signals over time and across multiple timescales (Fig. 1). Unlike traditional entropy measures that increase with degree of randomness, multiscale entropy is able to differentiate complex signals from white noise by considering entropy across multiple timescales. For example, Costa et al. (2005) compared multiscale entropy values for uncorrelated (white) noise versus correlated (pink) noise. While sample entropy

* Corresponding authors at: Rotman Research Institute at Baycrest, 3560 Bathurst Street, Toronto, Ontario, Canada M6A 2E1.
E-mail addresses: jheisz@rotman-baycrest.on.ca (J.J. Heisz), rmcintosh@rotman-baycrest.on.ca (A.R. McIntosh).

1053-8119/$ – see front matter.
http://dx.doi.org/10.1016/j.neuroimage.2012.08.018

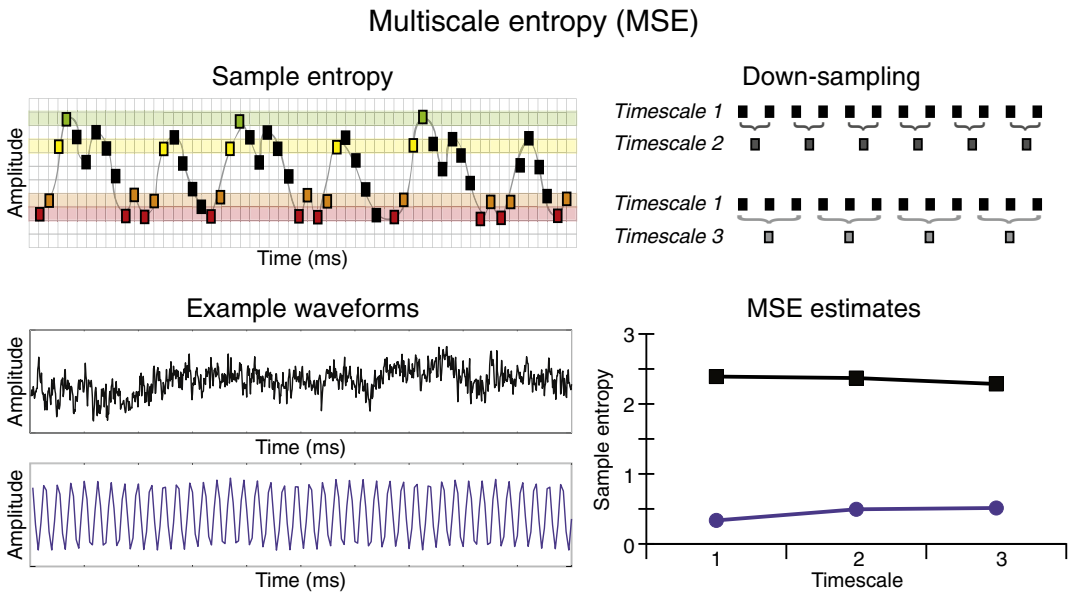


Fig. 1. Multiscale entropy (MSE) quantifies the variance and correlated properties of the brain signal over time providing a window into dynamic network reconfiguration. Sample entropy estimates the variability of a time series. In this example, *m* (the pattern length) is set to two, which means that the sequence pattern of two consecutive data points is considered; *r* (the similarity criterion), reflects the amplitude range (denoted by the height of the colored bands) within which data points are considered to "match". To calculate sample entropy for this simulated time series, begin with the first two data points. First, count the number of times two consecutive data points have amplitude values within the range of the red and orange color bands, respectively. There are 10 matches for this two-component sequence. Second, count the number of times three consecutive data points have amplitude values within the range of the red, orange and yellow color bands, respectively. There are 5 matches for this three-component sequence. Continue in this manner for the second and third data points in the time series. The number of two-component matches (5) and three-component matches (3) in this sequence is added to the previous values (total two-component matches = 15; total three-component matches = 8). Repeat for all data points in the time series (up to N–m) to determine the total ratio of two-component matches to three-component matches. Sample entropy is the natural logarithm of this ratio. A predictable waveform, depicted in purple, has a ratio near one and sample entropy near zero. A more variable waveform, depicted in black, has a ratio greater than 1 and sample entropy greater than zero. Down-sampling generates multiple time series of varying timescales. Timescale 1 is the original time series. To create the time series of subsequent timescales, simply divide the original time series into non-overlapping windows of the timescale length and average the data points within each window. Sample entropy is calculated for each timescale, hence multiscale entropy. Sample entropy and down-sampling illustrations were adopted from Costa et al. (2005).

was greater for white noise than pink noise at fine timescale, the opposite was observed at coarser timescales 5–20. In other words, when entropy was considered across multiple timescales, the true complexity of the signals was accurately represented. With respect to EEG signals, MSE is sensitive to the complexity of the oscillatory components contributing to the signal; this includes the number of components as well as the interactions between those components (i.e., nonlinear dynamics) at various frequencies (Vakorin and McIntosh, 2012). It is important to note that the present application of MSE to the stimulus-evoked response is relative in the sense that we are capturing changes in variability of the brain's response to a stimulus and not general intrinsic variability. These task-driven transient fluctuations in the oscillatory structure of the brain signal over time may reflect transitions or bifurcations between network microstates (Deco et al., 2011; Friston, 2001) driven by the face's familiarity. Accordingly, the variability of the signal as quantified by MSE may be used to estimate the degrees of freedom or complexity of the underlying network related to knowledge representation. Critically, two common methods of EEG analysis — mean amplitude and spectral power analysis, are not sensitive to the nonlinear stochastic activity.

Materials and methods

Experiment 1

Participants

Forty-two McMaster University undergraduate students participated. Half viewed non-famous faces and the other half viewed famous faces. Seven additional participants were run in the famous face condition to increase power for the correlation analysis. All participants were Caucasian and reported normal or corrected-to-normal vision. Participants provided written informed consent and received course credit for their participation. All procedures complied with the Canadian tri-council policy on ethics as approved by the McMaster Ethics Research Board.

Apparatus

Stimulus presentation and manual response measurement were controlled by Presentation experimental software (Version 11), running on a Pentium 4 Computer under Windows XP operating system. The stimuli were displayed on a 17-inch color CRT display at a resolution of 1280 × 1024 and frame refresh rate of 85 Hz. Participants were seated 80 cm from the display and the experiment was run in a dimly lit room.

Stimuli

The face stimuli consisted of 40 images of unfamiliar Caucasians (20 male) and 40 images of famous Caucasian celebrities (20 male) obtained via the worldwide web. All images captured the front of the face, with a neutral or smiling expression, without glasses. Images were cropped to include hair but exclude background, and converted to gray-scale. All images were resized to equate height, width, and resolution.

Procedure

The experiment consisted of 120 trials. The 40 faces (participants either viewed famous faces or unfamiliar faces) were randomly split into two equal sets, A and B. During the first 60 trials, the faces of set A were presented once and faces of set B were presented twice successively, with the order of presentation randomized. During the

last 60 trials, the faces of set B were presented once and the faces of set A were presented twice successively, with the order of presentation randomized. In total, each of the 40 faces was presented three times, once as a single and once as a repetition.

Each trial began with a fixation point presented for 1000 ms followed by a facial image presented until manual response. Participants were instructed to determine if the current face was the same as or different from the previous face by pressing 1 or 2 on the number pad for same and different responses, respectively (response buttons were counterbalanced across subjects). Both speed and accuracy were emphasized.

At the end of the experiment, participants rated the famous faces on a scale from 1 (not familiar) to 7 (highly familiar), and generated (if they could) the first and last name of the famous individual.

Experiment 2

Participants

Twenty McMaster University undergraduate students participated. All participants reported normal or corrected-to-normal vision. Eligible participants received course credit plus $15 for their participation, and the remainder received $40 compensation. Participants provided written informed consent. All procedures complied with the Canadian tri-council policy on ethics as approved by the McMaster Ethics Research Board.

Apparatus

The apparatus was the same as Experiment 1.

Stimuli

The face stimuli consisted of ten colored images of Caucasian individuals (five male) with neutral expressions. Faces were without glasses and a black wrap concealed clothing. Faces were adopted from a larger set provided by Dr. Daphne Maurer's Visual Developmental Lab, Department of Psychology, Neuroscience and Behaviour, McMaster University (Geldart et al., 1999). The images were presented at the center of the display on a gray background, approximately 5° of visual angle wide and 6° of visual angle high. Eight images (four males) were selected per subject and assigned to a particular condition; the assignment of face stimuli to particular conditions was counterbalanced across subjects.

Procedure

A detailed description of the stimuli and procedure can be found in Heisz and Shedden (2009). In brief, learning took place over five consecutive days with each session approximately 4.5 min. Across all sessions, two faces were presented for approximately 21 min each (high familiarity), and two faces were presented for approximately 3 min each (medium familiarity). An auditory story always accompanied face presentation and the story content either pertained to the face (i.e., details describing the individual's life events) or not (i.e., science articles). We conducted the multiscale entropy analysis for each of the story conditions separately; however, similar outcomes were observed and so the conditions were collapsed.

EEG was recorded before and after learning. Each EEG session lasted approximately 1 h. The pre-learning EEG session consisted of six face stimuli: four (to be) learned faces and two additional faces. The post-learning EEG session consisted of eight stimuli: the four learned faces (two high familiarity and two medium familiarity), the two faces from the pre-learning EEG session (low familiarity) plus two novel faces (novel). Faces were presented for 750 ms followed by an inter-stimulus interval of 750 ms, during which a fixation point was presented. Faces were presented in pseudo-randomized order; each face was presented approximately 200 times. Subjects performed a passive viewing task in which they were instructed to consider each face's identity.

Experiments 1 and 2

EEG acquisition and analysis

The ActiveTwo Biosemi electrode system was used to record continuous EEG activity from 128 Ag/AgCl scalp electrodes plus four additional electrodes placed at the outer canthi and just below each eye for recording horizontal and vertical eye movements. Two additional electrodes, common mode sense (CMS) active electrode and driven right leg (DRL) passive electrode were also used (http://www.biosemi.com/faq/cms&drl.htm). The continuous signal was acquired with an open pass-band from DC to 150 Hz and digitized at 512 Hz.

The analysis was conducted on 76 electrodes that corresponded with the 10–20 system. The continuous EEG signal was bandpass filtered between .1 Hz and 55 Hz and re-referenced to a common average reference. Data were epoched and baselined into epochs (1000 ms and 700 ms for E1 and E2, respectively) with a 100 ms pre-stimulus baseline. Preliminary artifact removal was performed using independent component analysis (ICA) as implemented in EEGLAB software (Delorme and Makeig, 2004). Trials contaminated with excessive amplitudes were removed first, then ICA decomposition was performed on the remaining concatenated trials and components carrying ocular and muscle artifacts were subtracted.

Multiscale entropy analysis

Multiscale entropy (MSE) was used to estimate variability at different timescales (Fig. 1; Costa et al., 2005, Heisz & McIntosh, in press). The utility of MSE has been confirmed by numerous studies (Bhattacharya et al., 2005; Catarino et al., 2011; Lippé et al., 2009; McIntosh et al., 2008; Misic et al., 2010; Mizuno et al., 2010). There is also a well-articulated theoretical framework that emphasizes the space-time structure as vital to understanding brain mechanisms of cognition and behavior (Breakspear and McIntosh, 2011; Deco et al., 2011); informatic theoretical measures like MSE are particularly sensitive to this space-time structure.

To calculate MSE, we used the algorithm available at www.physionet.org/physiotools/mse/, which computes MSE in two steps. First, the algorithm progressively down-samples the EEG post-stimulus time series $\{x_1,\ldots, x_i,\ldots, x_N\}$ per trial and per condition. For timescale τ, the coarse-grained time series $\{y^{(\tau)}\}$ is constructed by averaging data points within non-overlapping windows of length τ. Each element of a coarse-grained time series, j, is calculated according to Eq. (1):

$$y_j^{(\tau)} = \frac{1}{\tau}\sum_{i=(j-1)\tau+1}^{j\tau} x_i,\ 1 \le j \le \frac{N}{\tau}. \quad (1)$$

For example, timescale 1 is the original time series N (308 and 462 digitized data points for Experiment 1 and 2, respectively), which was down-sampled up to the tenth timescale (308/10 and 462/10 digitized data points for Experiment 1 and 2, respectively). To convert timescale into milliseconds, divide the timescale by the EEG sampling rate (512 Hz).

Second, the algorithm calculates the sample entropy for each coarse-grained time series (Eq. (2)):

$$S_E(m,r,N) = \ln\frac{\sum_{i=1}^{N-m} n_i^{'m}}{\sum_{i=1}^{N-m} n_i^{'m+1}}. \quad (2)$$

Sample entropy quantifies the variability by estimating the predictability of amplitude patterns across a time series of length N. The pattern length, m, was set to 2; that is, two consecutive data points were used for pattern matching. The similarity criterion, r, was set to .5; that is, data points were considered to have indistinguishable amplitude values (i.e., to "match") if the absolute amplitude difference between them was ≤50% of the time series standard

deviation. Note that $n_i^{'m}$ differs from n_i^m in that self-matches are not counted. For each subject, a channel specific MSE estimate was obtained as a mean across single trial entropy measures for timescales 1–10.

Spectral power

Spectral power density of each discrete time series x (i.e., for each subject, condition, channel and trial) was estimated by Welch's averaged modified periodogram method of spectral power estimation (Welch, 1967).

Statistical analysis

Statistical assessment of familiarity on MSE, spectral power, and mean amplitude scores was done using multivariate technique partial least squares (PLS) for EEG data (Krishnan et al., 2011; Lobaugh et al., 2001). PLS was performed on data matrices consisting of subject and channel specific measures such that rows represented subjects by condition. The columns of the data matrix contained the post-stimulus measures for MSE/spectral power/mean amplitude estimation by channel. The mean-centered matrices were decomposed with singular value decomposition (SVD) to identify the strongest condition differences and the corresponding scalp topography. This produced a set of orthogonal latent variables (LVs). Each LV consists of two parts: a "brain LV" (the brain portion of the LV) and a "design LV" (design portion of the LV). The brain LV represents the weighted linear combination of electrode sites and time points that co-vary with the design LV pattern. Projecting the brain LV onto each participant's EEG data by condition yields scalp scores, which can be positive or negative, depending on the relation between electrode/time (or electrode/frequency) and design LV. For brain-behavior analyses, correlations were computed between the questionnaire data and MSE/spectral power/mean amplitude measures across the entire sample.

The statistical significance of the effects was assessed using permutation tests (500) for the overall relationship between familiarity and brain response, and brain and behavior. The reliability of the topographies was determined with bootstrap estimation of confidence intervals, using 500 bootstrap samples. For scalp topographies, the singular vector weights for each channel were divided by the bootstrap estimated standard error, giving a bootstrap ratio. This is similar to a z score if the distribution of singular vector weights is Gaussian.

Results

Experiment 1

Does fame affect brain signal variability?

We assessed the relationship between knowledge representation and brain signal variability by measuring the MSE of EEG signals in response to stimuli that differ in the amount of stored information. Familiarity was manipulated by face fame. Participants either viewed famous faces or non-famous faces. Assuming that the famous faces were familiar to the participants, we expected the participants who viewed the famous faces to display a more variable brain response than those who viewed non-famous faces because familiar faces are associated with more information and have a richer memory network repertoire to dynamically explore. We also analyzed the same data set using the more traditional mean amplitude and spectral power approaches. If nonlinear stochastic processes support face familiarity then differences between familiar and unfamiliar faces should be revealed by MSE and not by the spectral power or mean amplitude methods of analysis.

While participants viewed the faces, they performed a 1-back identity-matching task in which they classified each face according to whether it was preceded by the same or different face. Although this was done to ensure that participants were attending to the face's identity, the task demands were expected to affect brain signal variability. In the present study, where the preceding and current face stimuli either matched perfectly or not at all, participants may have needed very little perceptual information to confirm a match between the preceding and the current face image. In other words, only partial reactivation of the network corresponding to the preceding face image may have been sufficient for an accurate "same" response. By this, we would expect to observe a decrease in sample entropy for repeated faces indicating a sparser network representation with less information processing capacity.

Task performance. Performance on the 1-back identity-matching task was faster but less accurate for repetitions (649 ms ± 17 ms; 96% ± 1%) compared to initial presentations (751 ms ± 28 ms; 98% ± 1%), suggesting a speed/accuracy trade-off consistent with the interpretation that the same responses were made using quick judgments based on little evidence. These observations were supported by repeated-measures analyses of variance conducted on mean response times and mean accuracy with a between-subjects factor of face fame (famous, non-famous) and a within-subjects factor of repetition (initial, repeat). Both analyses yielded a significant main effect of repetition [response time: $F(1,40) = 37.89$, $p < .001$, $ges = .11$; accuracy: $F(1,40) = 5.24$, $p < .05$, $ges = .05$]. There were no significant effects or interactions with face fame.

MSE. As illustrated by Fig. 2 (bar graph, left panel), the initial presentation of a face elicited more sample entropy than the immediate repetition of the same face ($p < .001$). Topographic plots depict the spatiotemporal distribution of this contrast. Although statistically reliable across all timescales, this contrast was more pronounced at the finer timescales. The spatial distribution of the contrast was particularly stable at electrodes over central parietal and prefrontal cortices. The contrast between famous and non-famous faces was not significant.

Spectral power and mean amplitude. As a comparison, the same analysis was conducted on spectral power and mean amplitude. Like MSE, spectral power and mean amplitude were sensitive to face repetition. For spectral power, initial presentations elicited greater power than repetitions for both famous and non-famous faces ($p < .001$; Fig. 2, middle panel). This repetition effect was broadly distributed over central parietal and prefrontal cortices and most stable at frequencies between 4 and 12 Hz. Frequencies less than 4 Hz had a similar topographical distribution but for the reverse contrast of greater power for repetitions than initial presentations.

For mean amplitude, a more negative response was elicited for repetitions than initial presentations for both famous and non-famous faces ($p < .001$; Fig. 2, right panel). This repetition effect was most stable at the N250 ERP component over bilateral occipital–temporal cortex, which is consistent with the typical response pattern observed for this component to repeated faces (Schweinberger et al., 2004). The reverse contrast (i.e., a more positive amplitude for repeat than initial presentation) was observed at the P3 ERP component over parietal cortex and at P1 and N170 ERP components over bilateral occipital–temporal cortex. This is also consistent with previous ERP studies on face repetition (Itier and Taylor, 2002).

Does familiarity with a famous face affect brain signal variability?

As part of the same experiment, we assessed the relationship between familiarity with the famous faces and brain signal variability. We correlated participants' behavioral reports of famous face familiarity with their MSE measures for those same famous faces. Behavioral reports were acquired at the end of the experiment; participants estimated their familiarity with the famous faces on a scale from 1 (not familiar) to 7 (extremely familiar) and recalled (if they could) the first and last name of the famous individual. We expected that

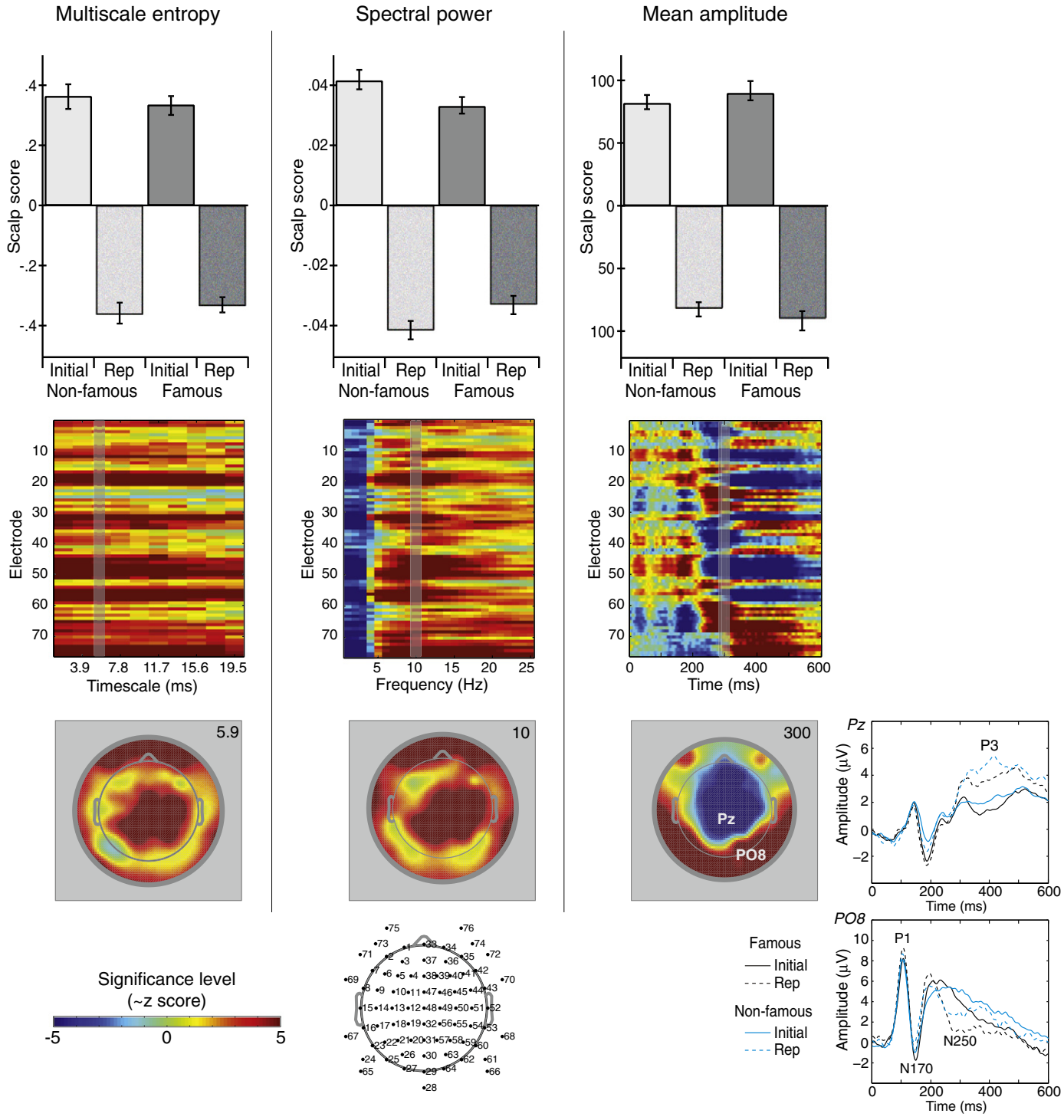


Fig. 2. Contrasting the EEG response to famous versus non-famous faces across measures of multiscale entropy, spectral power, and mean amplitude. The bar graphs (with standard error bars) depict contrasts between conditions that were significantly expressed across the entire data set as determined by permutation tests and bootstrap estimated 95% confidence intervals. The image plot highlights the electrodes and time/frequency at which the contrast across all conditions was most stable as determined by bootstrapping. Head plots illustrate the spatial distribution for a representative time/frequency. Values represent *~z* scores and negative values denote significance for the inverse condition effect.

participants who were more familiar with the famous faces would elicit a more variable brain response to those faces. Fig. 3 depicts the results.

Behavioral reports. Participants varied in their familiarity with the set of famous faces. The mean proportion of accurately named famous faces was .76 ± .23 SD, ranging from .2 to 1, which is significantly less than perfect naming accuracy of 1 ($t(27) = -5.74$, $p < .001$). Likewise, the mean familiarity rating was 5.6 ± .9 SD, ranging from 3.2 to 7, which is significantly less than the maximum familiarity rating of 7 ($t(27) = -7.88$, $p < .001$).

MSE. Brain signal entropy was higher for participants who were more familiar with the famous faces. Mean sample entropy values for the famous faces correlated with familiarity ratings and naming accuracy ($p < .05$; Fig. 3). Both rating and naming correlations were reliably

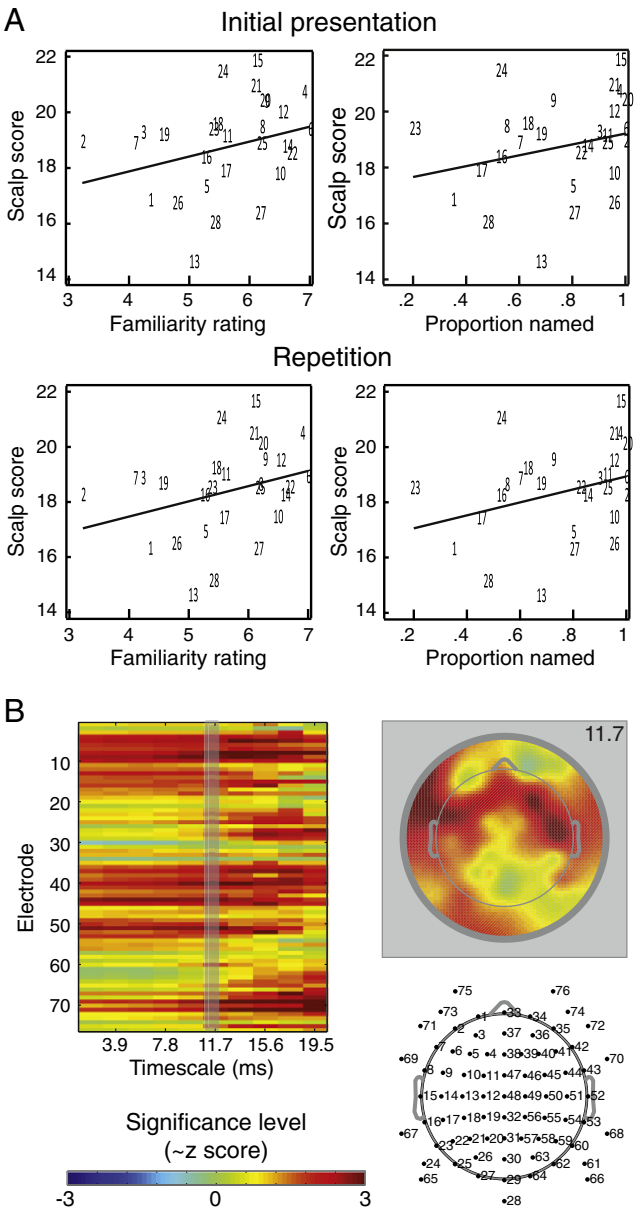


Fig. 3. Correlating multiscale entropy of the EEG response with behavioral measures of familiarity (ratings and namability) for famous faces. The scatter plots depict correlation between scalp scores and the behavioral responses for the famous faces that were significantly expressed across the entire data set as determined by permutation tests and bootstrap estimated 95% confidence intervals. Each number in the scatter plot represents a particular participant. The image plot highlights the electrodes and time at which the contrast across all conditions was most stable as determined by bootstrapping. Head plots illustrate the spatial distribution for a representative timescale. Values represent ~*z* scores and negative values denote significance for the inverse condition effect. Note: these behavioral measures of familiarity did not correlate with spectral power or mean amplitude.

expressed across all timescales over bilateral temporal–parietal extending up to frontal cortices.

Spectral power and mean amplitude. Spectral power and mean amplitude did not significantly correlate with familiarity measures of proportion named and familiarity ratings (all $ps > .05$).

Experiment 2

Does brain signal variability increase with learning?

In Experiment 2, we assessed whether newly acquired information builds up the variability in the brain signal. Rather than using famous and non-famous faces, participants were familiarized with previously novel faces to varying degrees over five consecutive days (Heisz and Shedden, 2009). Faces were familiarized to low, medium, or high levels; we did this by systematically increasing the amount of exposure the participant has with each face. Faces were also paired with either vignettes or science stories, and this contributed to the total amount of information gained. Preliminary analyses revealed similar outcomes for both learning conditions regardless of the type of paired information and so the factor of information type was collapsed. We recorded EEG before and after learning while participants passively viewed the faces and calculated the corresponding MSE per face and per session. We expected MSE values to increase as a function of face learning. Fig. 4 depicts the results.

MSE. Before learning, all faces were equally novel and, as expected, all faces elicited an equivalent MSE response. After learning, MSE values related to the amount of acquired familiarity such that faces seen more often during the experiment elicited a more variable response ($p < .001$; Fig. 4, left panel). The topographic plot of Fig. 4 (left panel) depicts the spatiotemporal distribution of this contrast. At coarse timescales, the acquired familiarity effect was most statistically reliable over right occipital–temporal cortex whereas at finer timescale the same effect showed a more anterior distribution.

Spectral power and mean amplitude. Like MSE, spectral power related to the amount of acquired familiarity ($p < .05$; Fig. 4, middle panel); however, unlike MSE, spectral power did not accurately distinguish medium and low familiarity faces according to the amount of prior exposure. The topographic plots of Fig. 4 (middle panel) depict the spatiotemporal distribution of this contrast. At lower frequencies between (5–12 Hz), the acquired familiarity effect was most statistically reliable over right occipital–temporal cortex whereas at higher frequencies the same effect showed a more anterior distribution. The reverse contrast was observed at frequencies less than 4 Hz.

Mean amplitude distinguished new faces from previously viewed faces but not among the previously viewed faces that varied in amount of prior exposure ($p < .001$; Fig. 4, right panel). Moreover, the effect of familiarity on the mean amplitude depended on scalp location. From 450 to 600 ms over parietal cortex new faces elicited a more positive response than previously viewed faces. During the same time window but over pre-frontal cortex the amplitude pattern was reversed. These results are consistent with previous ERP studies on stimulus novelty (e.g., Ferrari et al., 2010).

Discussion

We used MSE to measure brain signal variability while manipulating the amount of information associated with faces stored in memory. Individuals with greater familiarity for a group of famous faces elicited a more variable brain response. Brain signal variability also increased with learning after many exposures to previously unfamiliar faces. These results demonstrate an important relationship between brain signal variability and familiarity. Specifically, the results suggest that cognitive processes during the perception of familiar stimuli may engage a broader network of regions, which manifests as higher complexity/variability in spatial and temporal domains.

The measures of famous face familiarity and acquired familiarity correlated with MSE across all timescales. Previous research has demonstrated a direct link between MSE temporal scale and the bias between local versus distributed information. Local information is represented at finer timescales and distributed information is represented at coarser timescales (Vakorin et al., 2011). The present findings suggest that familiarity increases information processing capacity across both local and distributed regions. This makes sense if we consider that face familiarity/learning involves the buildup of information about the facial structure (local information) as well as

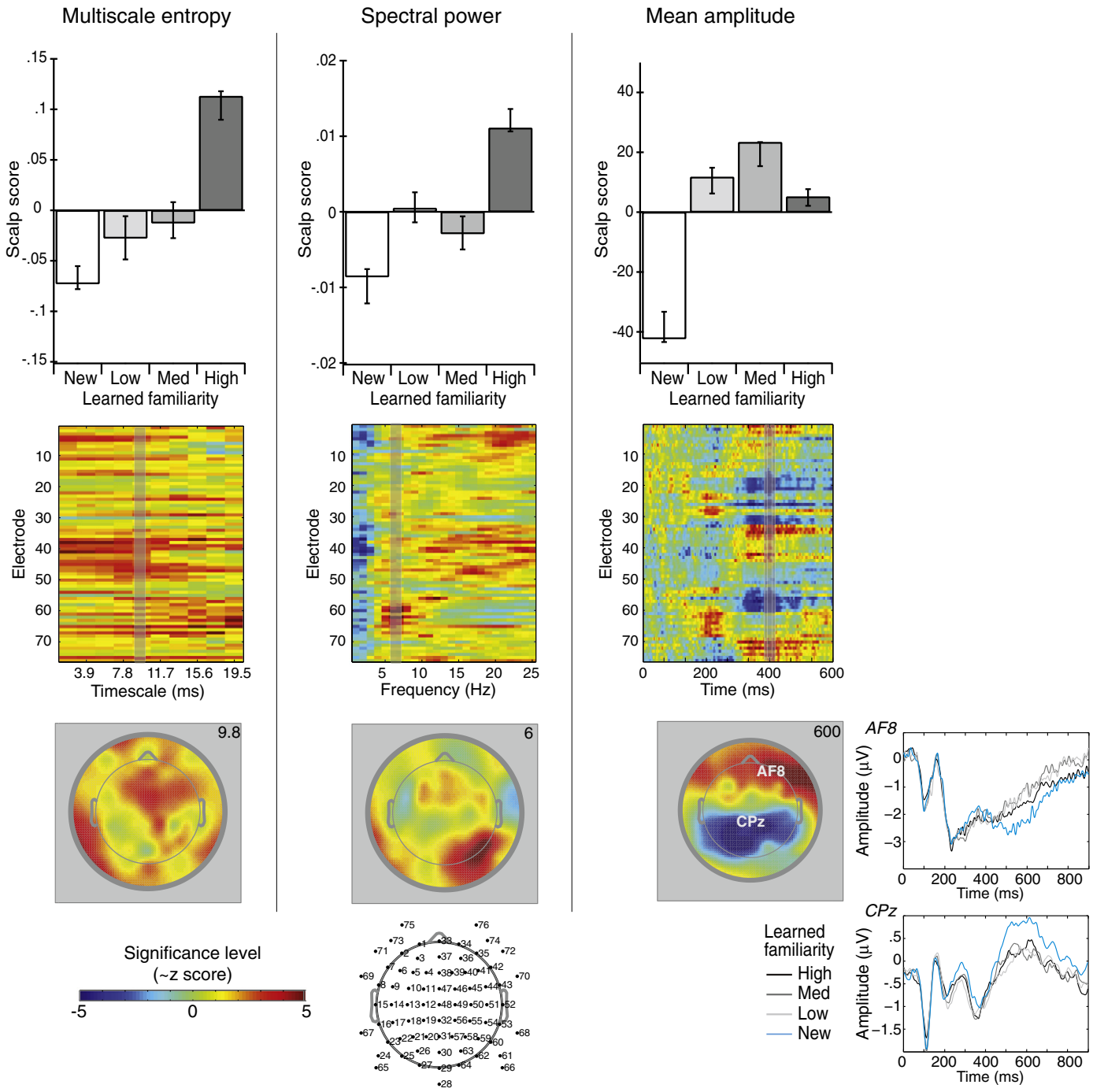


Fig. 4. Contrasting the EEG response to learned faces across measures of multiscale entropy, spectral power, and mean amplitude. The bar graphs (with standard error bars) depict contrasts between conditions that were significantly expressed across the entire data set as determined by permutation tests and bootstrap estimated 95% confidence intervals. The image plot highlights the electrodes and time/frequency at which the contrast across all conditions was most stable as determined by bootstrapping. Head plots illustrate the spatial distribution for a representative timescale. Values represent *~z* scores and negative values denote significance for the inverse condition effect.

linking that structure with associated information (distributed information). The information content (i.e., life events related to the faces versus unrelated science articles) paired with the face during learning did not seem to affect entropy values; however in this case all faces were learned with information, within the same laboratory setting, and over the same number of days; that is, there was an increase in information for all learned faces and this was accompanied by an increase in MSE.

In contrast, significant differences were not observed when contrasting famous versus non-famous faces. This is somewhat surprising given the stable fame effects typically observed in ERP (Begleiter et al., 1995; Trautner et al., 2004). Two things may be at play here. First, the contrast between famous and non-famous faces is usually done within subject and we used a between-subjects design. Second, and perhaps more importantly, the famous faces were not equally familiar to all the participants. In fact, some participants reported having very little familiarity with this particular group of famous faces (see Fig. 3 for familiarity rating and naming performance). Critically, this reduces the effect of familiarity that famous faces could have on the groups' brain response and minimizes brain response differences between famous and non-famous faces. As a consequence, the familiarity correlations of Experiment 1 and the acquired

familiarity comparison of Experiment 2 provide more robust estimates of the effect of familiarity on brain responses.

Overall, the analysis of EEG by MSE produced unique information that was not obtained using more traditional analyses, likely because the additional information gain by MSE reflects nonlinear network dynamics (Breakspear, 2002; Freeman and Rogers, 2002; Friston, 2001; McIntosh et al., 2008). Neither spectral power nor mean amplitude differentiated individual differences in famous face familiarity whereas MSE did. Furthermore, MSE was more sensitive to the subtle differences among experimentally learned faces than spectral power or mean amplitude. All measures converged for the effects of immediate stimulus repetition (Fig. 2), where repetition caused a decrease in both the mean and the variability of the brain's response; however, the observed decrease in entropy is important as it constrains the interpretation of potential neural processes that could produce the change in the mean response. Three putative models have been put forth (see Grill-Spector et al., 2006, for a review). According to the Fatigue model (Grill-Spector and Malach, 2001; Miller and Desimone, 1994), repetition suppression reflects a decrease in mean population firing rate with no change in the relative response of the network nodes. According to the Sharpening model (Desimone, 1996; Wiggs and Martin, 1998), repetition suppression reflects the activity of fewer, more selective neurons resulting in an overall reduction in the number of network nodes that respond to the repeated stimulus. According to the Facilitation model (Henson and Rugg, 2003), repetition suppression does not reflect a change in the overall activation pattern but rather reflects faster processing across the same network. Although all models could be used to explain the observed drop in mean amplitude with stimulus repetition, only the Sharpening model predicts a drop in entropy. Specifically, the Sharpening model proposes sparser network coding for repeated stimuli, which implies fewer potential configurations and thus a less variable response (Deco et al., 2011; Ghosh et al., 2008; Tsuda, 2001).

We propose that brain signal variability, as measured by MSE, may reveal the information processing capacity of the system and provide an index of the size of the repertoire of responses produced by the brain when a stimulus is presented. MSE provides an index of network complexity that cannot be obtained by simply counting the number of active brain regions; although the characterization of network architecture is important, it represents the static picture that is blind to transient, dynamic interactions between network nodes. Upon stimulus presentation, information stored in associated networks becomes active and integrated through dynamic network reconfiguration (Deco et al., 2011; Tononi et al., 1998). This may affect the variance of local signaling by continuously altering the relative weights of reentrant inputs from active feedback and lateral connections (Fuster, 1997). Ultimately, the number of oscillatory components contributing to the EEG signal as well as the interactions between those components would change over time and across timescales, and this would be reflected in the MSE values of the corresponding time series (Vakorin and McIntosh, 2012). Indeed, previous work has used MSE to reveal long-range power-law correlations within the spontaneous firing pattern of neurons in human hippocampal and amygdala structures (Bhattacharya et al., 2005); critically, these characteristics promote the flexibility that a system needs to explore its dynamic repertoire (Deco et al., 2011; Ghosh et al., 2008). It follows that the more information (repertoire) available for a given stimulus means the more signal variability elicited by that stimulus (Deco et al., 2011). Here, familiar faces produced greater variability than unfamiliar faces. Since familiar faces are associated with more information than unfamiliar faces (Burton et al., 1999), our results support the hypothesis that signal variability is an expression of the exploration of the dynamic repertoire of the system.

The proposed relationship between network information processing capacity and signal variability fits nicely within the theoretical framework of complexity matching (Tononi et al., 1996, 1998). Through computer simulations, Tononi et al. (1996) demonstrated that the amount of information available for a stimulus could be determined by the extent to which the complexity of the stimulus matched the complexity of its corresponding brain response. A stronger match was observed for familiar stimuli compared to novel stimuli, suggesting less information available for the processing of novel stimuli. In the present study, the differential effects of stimulus novelty on signal mean versus signal variability may be related to complexity matching. Compared to familiar faces, novel faces elicited a larger evoked response over parietal cortex (Fig. 4, right panel) and the signal was less variable. It is possible that the evoked response captures the restructuring of the underlying correlation weights of the network architecture in order to establish “meaning” or context for the novel stimuli. As information builds up in the network's correlational properties, the repertoire of brain responses associated with that stimulus increases resulting in a more variable brain signal.

Our results are counter to the notion of memories as stable associative attractors (e.g., Hopfield, 1982). In the context of an associative attractor network, familiar stimuli elicit unique connectivity patterns, which are represented by a single attractor state. Novel stimuli are not associated with a particular attractor state and this causes the system to jump between multiple possible matches eliciting a more variable response. A critical feature of the associative attractor network is that once an attractor state is achieved the system remains fixed in that state. The present results do not refute the existence of associative attractors or their involvement in memory representation; however, the observation that familiar stimuli elicit a more variable response than unfamiliar stimuli suggests that, during memory retrieval, network activity is not fixed but fluid with continuous switches between configurations.

Conclusion

In conclusion, the results converge to reveal a robust relationship between memory content, brain signal variability, and the corresponding network dynamics. Our novel contribution illuminates the importance of brain signal variability, demonstrating that variability during memory processing provides a window into knowledge representation. Highly familiar faces associated with greater information content elicited a more variable response. Including variability as a key neurological marker of network information processing capacity provides critical information about the complexity of the underlying neural network dynamics and the richness of the associated memory. The present observations, together with the existing empirical and theoretical work exploring the properties of neural network dynamics (Deco et al., 2011; Tsuda, 2001; Varela et al., 2001), move us away from the characterization of mental function as “states” and towards a fluid unfolding of processes that link to human cognition.

Acknowledgments

This research was supported by the James S. McDonnell Foundation grant awarded to ARM and the Natural Sciences and Engineering Research Council of Canada (NSERC) grant awarded to JMS.

References

Begleiter, H., Porjesz, B., Wang, W., 1995. Event-related brain potentials differentiate priming and recognition to familiar and unfamiliar faces. Electroencephalogr. Clin. Neurophysiol. 94, 41–49.
Bhattacharya, J., Edwards, J., Mamelak, A.N., Schuman, E.M., 2005. Long-range temporal correlations in the spontaneous spiking of neurons in the hippocampal and amygdala complex of humans. Neuroscience 131, 547–555.
Breakspear, M., 2002. Nonlinear phase desynchronization in human electroencephalographic data. Hum. Brain Mapp. 15, 175–198.
Breakspear, M., McIntosh, A.R., 2011. Networks, noise, and models: reconceptualizing the brain as a complex, distributed system. NeuroImage 58, 293–295.
Burton, A.M., Bruce, V., Hancock, P.J.B., 1999. From pixels to people: a model of familiar face recognition. Cogn. Sci. 23, 1–31.

Catarino, A., Churches, O., Baron-Cohen, S., Andrade, A., Ring, H., 2011. Atypical EEG complexity in autism spectrum conditions: a multiscale entropy analysis. Clin. Neurophysiol. 122, 2375–2383.
Costa, M., Goldberger, A., Peng, C., 2005. Multiscale entropy analysis of biological signals. Phys. Rev. E 712, 1–18.
Deco, G., Jirsa, V., McIntosh, A.R., 2011. Emerging concepts for the dynamical organization of resting-state activity in the brain. Nat. Rev. Neurosci. 12, 43–56.
Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Meth. 134, 9–21.
Desimone, R., 1996. Neural mechanisms for visual memory and their role in attention. PNAS 93, 13494–13499.
Ferrari, V., Bradley, M.M., Codispoti, M., Lang, P.L., 2010. Detecting novelty and significance. J. Cogn. Neurosci. 22, 404–411.
Freeman, W.J., Rogers, L.J., 2002. Fine temporal resolution of analytic phase reveals episodic synchronization by state transitions in gamma EEGs. J. Neurophysiol. 87, 937–945.
Friston, K.J., 2001. The labile brain. I. Neuronal transients and nonlinear coupling. Philos. Trans. R. Soc. Lond. B Biol. Sci. 355B, 215–236.
Fuster, J.M., 1997. Network memory. Trends Neurosci. 20, 451–459.
Garrett, D.D., Kovacevic, N., McIntosh, A.R., Grady, C.L., 2011. The importance of being variable. J. Neurosci. 31, 4496–4503.
Gatlin, L., 1972. Information Theory and the Living System. Columbia University Press, New York.
Geldart, S., Maurer, D., Henderson, H., 1999. Effects of the height of internal features of faces on adults' aesthetic ratings and 5-month-olds' looking time. Perception 28, 839–850.
Ghosh, A., Rho, Y., McIntosh, A.R., Kotter, R., Jirsa, V.K., 2008. Noise during rest enables the exploration of the brain's dynamic repertoire. PLoS Comput. Biol. 4 (10), e1000196.
Grill-Spector, K., Malach, R., 2001. fMR-adaptation: a tool for studying functional properties of human cortical neuron. Acta Psychol. 107, 293–321.
Grill-Spector, K., Henson, R., Martin, A., 2006. Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn. Sci. 10, 14–23.
Haxby, J.V., Hoffman, E.A., Gobbini, M.I., 2000. The distributed human neural system for face perception. Trends Cogn. Sci. 4, 223–233.
Heisz, J., McIntosh, A. R., (in press). Applications Of EEG Neuroimaging Data: Event-related Potentials, Spectral Power, And Multiscale Entropy. J. Vis. Exp. e50131, http://dx.doi.org/10.3791/50131
Heisz, J.J., Shedden, J.M., 2009. Semantic learning modifies perceptual face processing. J. Cogn. Neurosci. 216, 1127–1134.
Henson, R.N., Rugg, M.D., 2003. Neural response suppression, haemodynamic repetition effects and behavioural priming. Neuropsychologia 41, 263–270.
Honey, C., Kotter, R., Breakspear, M., Sporns, O., 2007. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. PNAS 104, 10240–10245.
Hopfield, J.J., 1982. Neural networks and physical systems with emergent collective computation abilities. PNAS 79, 2554–2558.
Itier, R.J., Taylor, M.J., 2002. Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: a repetition study using ERPs. NeuroImage 15, 353–372.
Krishnan, A., Williams, L.J., McIntosh, A.R., Abdi, H., 2011. Partial least squares (PLS) methods for neuroimaging: a tutorial and review. NeuroImage 56 (2), 455–475.
Lippé, S., Kovacevic, N., McIntosh, A.R., 2009. Differential maturation of brain signal complexity in the human auditory and visual system. Front. Hum. Neurosci. 3, 48.
Lobaugh, N.J., West, R., McIntosh, A.R., 2001. Spatiotemporal analysis of experimental differences in event-related potential data with partial least squares. Psychophysiology 38, 517–530.
McIntosh, A.R., Kovacevic, N., Itier, R.J., 2008. Increased brain signal variability accompanies lower behavioral variability in development. PLoS Comput. Biol. 47, e1000106.
Miller, E.K., Desimone, R., 1994. Parallel neuronal mechanism for short-term memory. Science 263, 520–522.
Misic, B., Mills, T., Taylor, M.J., McIntosh, A.R., 2010. Brain noise is task-dependent and region-specific. J. Neurophysiol. 104, 2667–2676.
Mizuno, T., Takahashi, T., Cho, R.Y., Kikuchi, M., Takahashi, K., Wada, Y., 2010. Assessment of EEG dynamical complexity in Alzheimer's disease using multiscale entropy. Clin. Neurphysiol. 121, 1438–1446.
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.
Schweinberger, S.R., Huddy, V., Burton, A.M., 2004. N250r: a face selective brain response to stimulus repetition. Neuroreport 15, 1501–1505.
Shannon, C.E., 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27, 379–423.
Tononi, G., Sporns, O., Edelman, G.M., 1996. A complexity measure for selective matching of signals by the brain. PNAS 93, 3422–3427.
Tononi, G., Edelman, G.M., Sporns, O., 1998. Complexity and coherency: integrating information in the brain. Trends Cogn. Sci. 212, 474–484.
Trautner, P., Dietl, T., Staedtgen, M., Mecklinger, A., Grunwald, T., Elger, C.E., Kurthen, M., 2004. Recognition of famous faces in the medial temporal lobe: an invasive ERP study. Neurology 63, 1203–1208.
Tsuda, I., 2001. Towards an interpretation of dynamic neural activity in terms of chaotic dynamical systems. Behav. Brain Sci. 24, 826–827.
Vakorin, V.A., McIntosh, A.R., 2012. Mapping the multi-scale information content of compelx brain signals, in Principles of brain dynamics: global state interactions. In: Rabinovich, M., Friston, K., Varona, P. (Eds.), MIT Press.
Vakorin, V.A., Lippé, S., McIntosh, A.R., 2011. Variability of brain signal processed locally transforms into higher connectivity with brain development. J. Neurosci. 31, 6405–6413.
Varela, F., Lachaux, J.P., Rodriguez, E., Martinerie, J., 2001. The brainweb: phase synchronization and largescale integration. Nat. Rev. Neurosci. 2, 229–239.
Welch, P.D., 1967. The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE Trans. Audio Electroacoust. AU-15, 70–73.
Wiggs, C.L., Martin, A., 1998. Properties and mechanism of perceptual priming. Curr. Opin. Neurobiol. 8, 227–233.