



OPEN Brain complexity represents uncertainty in sequence learning and corroborates habituation deficit in Parkinson disease patients

Mohammad Hossein Heidari Beni, Kamyab Hosseinpour, Mohammad Reza Seyednejad & Hamid Aghajan✉

Information processing is a core element of cognitive processes in the brain. Rooted in the idea postulated by John von Neumann that the stochasticity of neural operations might essentially be a reflection of the probabilistic nature of information, it can be argued that fluctuations in the complexity of the brain operations should reflect aspects of information processing activities undertaken to infer nuances embedded in the sensory data. The current study elucidates these aspects by examining the intricate relationship between information processing in the brain, reflected in the complexity of brain activities, and information-theoretical measures of uncertainty in the Bayesian models of sequence learning. A key finding of our study is that the brain complexity shows a significantly superior correlation with the uncertainty about the input, captured by the entropy of the posterior distribution over all input values, compared to its correlation with the novel information stored in the current input as measured by the Shannon or Bayesian models of surprise. Additionally, we found that the brain complexity follows the habituation process known to occur during learning by demonstrating a decreasing trend and also corroborates the process of spontaneous recovery at the onset of a repeated learning task. As further evidence for this relationship, distinguished complexity trends are observed in patients of Parkinson disease, reflecting cognitive disorders in habituation and habit formation caused by the disease.

John von Neumann predicted some 65 years ago that eventually both the brain theory and the theory of computation would come to resemble the physics of statistical mechanics and thermodynamics¹. His intuition was based on the stochastic character of neural activity in the brain and hence the necessity of resorting to statistical models for explaining various aspects of neural interactions¹. Based on this assessment, and as evidenced by the subsequent decades of efforts to utilize tools of information theory to describe neural operations, an even more illuminating insight into the stochasticity of neural operations stems from the perspective that they might essentially be a reflection of the probabilistic nature of information¹.

On a similar line of analogy, one can argue that fluctuations in the complexity of the operations which the brain undertakes should reflect the nuances in the surrounding environment embedded in the collected sensory information. In other words, the stochasticity in neural activities of the brain can be interpreted as a necessary attribute of forming a response to the uncertainty sensed in the input data. It can be hence argued that uncertainty about the data should be mirrored in the stochastic characteristics of brain activity. In computational terms, this relationship can be explained as a manifest of the brain's allocation of more resources to incorporate the nuances observed in the sensory data into its model of the environment. This means that the more uncertain we are about our surroundings, the more our neural activities ramp up information processing to extract useful information. As we become more certain about our environment over time, the level of information processing decreases as the demand for extracting more information is reduced. The intensity of information processing is reflected in the complexity of brain activities over time and space².

While the relationship between the input information content and the brain's response has long been studied³, there is little information on how the brain's operational complexity and its fluctuations in time are shaped by

Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran. ✉email: aghajan@ee.sharif.edu

the statistical characteristics of the surrounding sensory environment from which information flows into and propagates throughout the brain. Many studies have highlighted the importance of complexity in the brain in performing cognitive tasks^{2,4,5}. However, there is a paucity of reports on the different aspects of learning and their relationship with the complexity within the brain. One potential link that can bring both learning (as a high-level overarching purpose) and complexity (as a substrate-level information handling activity) within a common framework is that both have strong connections with information processing involved in a cognitive process.

Interactions between different neural circuits and their dynamics over time are known to play a critical role in the integration of information from different sensory sources and memory which leads to the execution of advanced computations that contribute to cognitive abilities^{6,7}. Given the stochastic nature of these interactions, the complexity of the brain circuitry can potentially be influenced by nuances of the network dynamics during a cognitive process. In this premise, the complexity of the brain can primarily serve as an indicator of the form and structure of network-level interactions, which reflect the dynamics of the information flow and can be described not only in spatial terms but also over time.

Thus, the main dimensions involved in the complexity of the brain can be identified as connectivity (spatial), dynamics (temporal), and information (external drive)⁴. Connectivity can be categorized into two forms: structural and functional⁸, both driven by the spatial interaction among brain regions, and thus offering a framework for understanding information processing across the brain as an essential operation in cognitive tasks^{9,10}. A plethora of information-theoretical measures of spatial interaction has been introduced as brain complexity measures and applied to recorded neural activity or EEG data². The changes observed in such complexity measures during a cognitive task can be interpreted as an indication of variations in the functional connectivity of the brain, which are manifests of the establishment of cognitive controls and execution of cognitive tasks.

In addition to the spatial interactions across brain regions which are often modeled in a stochastic framework through information-theoretical measures such as mutual information², other inter-dependencies which manifest in the generation and integration of information over time can be modeled via measures of information nuances used in time series analysis. In this context, the temporal dynamics of the measured brain activity can be attributed to the dynamics of computational states that form cognitive functions such as learning. Several complexity measures have been in use for time series data, defined to quantify the rate of information generation over time. The Lempel–Ziv complexity¹¹ and sample entropy¹² operate on time series and approximate the entropy rate for stochastic processes^{11,13}. To summarize, complexity, as defined in an information-theoretic framework, can be regarded as an indication of information integration among brain regions and over time, influenced respectively, by the connectivity of the brain and the temporal dynamics of brain activity.

Numerous studies have explored the relationship between entropy and complexity as information-related measures, both mathematically^{14–17} and in complex dynamical systems^{18–20}. For instance¹⁴, has shown that complexity is the time derivative of entropy. Moreover, in the context of modeling the activities of the brain as a complex system, the interrelationship between complexity and entropy has been explored both in response to input stimuli²¹ and for different brain functions^{22,23}, albeit primarily in studies of consciousness^{23–25}, rather than in models of learning. While entropy represents encoded information²⁶, complexity reflects information integration². In this study, we aim to develop a computational, yet intuitive relationship between the complexity of brain activities and entropy as uncertainty in Bayesian learning. Indeed, we posit that, in addition to the known quantitative relationship between the entropy of an input sequence and its complexity^{16,17}, entropy as a measure of uncertainty about the input data in the context of Bayesian learning²⁷ correlates with the complexity of brain activities as both reflect the need to increase the level of information integration to reduce uncertainty through a learning process.

In a general view, a learning process can be mathematically interpreted as the process of finding an optimal conditional distribution for actions and perception, based on the received observations. While such conditional distributions can be formed during a range of activities from encoding task-related information to making predictions and decisions guided by goal-directed behaviors, one fundamental type of learning occurs when receiving input sequences during which most observations reported by sensory systems exhibit a sequential structure. Therefore, understanding the learning mechanisms for sequences can provide insight and help establish a foundation for studying other types of learning such as language acquisition²⁸ and adaptive behavior²⁹. Intuitively, learning a sequence involves inferring the procedure for the generation of each element in the sequence. Mathematically, this entails finding the conditional distribution of the upcoming stimulus given the set of previous stimuli. Assuming the Markov property³, proposes a local transition probability model for human sequence learning, linking statistical features to behavior (e.g. reaction time) and brain activity (e.g. P300 amplitude). Through analyzing surprise in input sequences³⁰, demonstrates how a local transition probability model can predict the effects of local and global statistics on brain's response in alternation-biased and repetition-biased sequences, and how these statistics, encoded at varying temporal scales, shape neural activity. Additionally, brain regions exhibit distinct temporal coding patterns based on stimuli or choices, which play key roles in shaping behavior^{31–33}. Therefore, given the significant role of temporal patterns in information processing, studying the relationship between the entropy rate of the brain activity dynamics and the statistics in transition probability model may offer further insight to elucidate computational processes underlying sequence learning.

One of the fundamental processes in sequence learning involves integrating information generated by the current stimulus with information gathered from previous stimuli which may be stored in memory. This integration can be modeled by forming a posterior distribution on the next stimulus, mathematically known as Bayesian inference^{3,34–36}. A central characteristic of information processing in sequence learning is the issue of handling surprise. Two notions of surprise are often derived in modeling the process of sequence learning:

puzzlement surprise and enlightenment surprise. Puzzlement surprise is linked to the level of conflict between the internal model and the world after observing a new stimulus³⁴ and is mainly quantified by Shannon surprise, which regards the surprise associated with a stimulus as its unlikeliness. Enlightenment surprise is related to the amount of change in the internal model after observing a new stimulus³⁴, and is also termed as Bayesian surprise since it is often formulated as the Kullback–Leibler (KL) divergence between the prior distribution of the model parameters and the posterior distribution calculated after observing a new stimulus.

On the other hand, uncertainty about the next stimulus, measured by the entropy of the conditional distribution of the next stimulus given the previous samples of stimuli (i.e. the same posterior distribution), is closely related to the amount of information garnered when the next stimulus is revealed. Moreover, in the context of a dynamical system, information integration relates to the rate of information generated as the dynamical states of the system evolve, or to the reduction in uncertainty as the dynamics transition to the next state³⁷. Therefore, one potential area to investigate the implications of information integration in the brain and the control mechanisms for handling uncertainty, which are essential in sequence learning, is the complexity of the dynamics of brain activity. Furthermore, according to the free energy principle³⁸, the brain continuously attempts to minimize uncertainty, which can be quantified as the entropy or the average surprise value across all possible random states. Neural representations, particularly at higher levels beyond sensory processing, are formed through perception and encode the causes of sensory data through bottom-up processes²⁷. Uncertainty about the causes of sensory data, known as world states, is thus manifested in the complexity of neural activities. This uncertainty is propagated through top-down processes and contributes to the generation of prediction error signals^{27,38,39}. In the context of sequence learning, uncertainty about the next stimulus, captured by the entropy of the conditional distribution learned up to that point, can be potentially reflected in the complexity of neural activities during the observation of a sequence of stimuli. This hypothesis can be validated by investigating the relation between the entropy of the learned conditional distribution and the complexity of brain activities recorded during the observation of the input sensory sequence.

Attention and habituation serve as underlying computational activities of the brain related to the process of learning. From a computational perspective, attention is a mechanism that aims to minimize uncertainty during the learning process, which is reflected in statistical aspects of the learned conditional distribution, such as its variance or entropy²⁷. On the other hand, habituation can be seen as a learning process in which the updating rate of the conditional distribution decreases, leading the learner to converge to an optimal conditional distribution. In terms of brain activities, attention manifests itself through the uncertainty level or the entropy of neural activities²⁷. Consequently, habituation is characterized by a form of attenuation in brain activity properties^{40,41}, which can be associated with a reduction in the level of attention.

In addition to the aforementioned properties associated with habituation, linking habituation to a decrease in the complexity of brain activities during the learning process could offer a more consistent and insightful understanding of information integration from a computational standpoint. Intuitively, in habituation, the process of forming a behavioral conditional distribution occurs under a diminishing pursuit of information integration in shaping the form of the conditional distribution. Considering the relationship between information integration, attention, and complexity^{2,42–44} as outlined above, this multi-faceted perspective provides further rationale for why habitual behaviors, resulting from the habituation process, occur with a lower level of conscious awareness or attention⁴⁵.

Parkinson disease (PD) is a common progressive neurodegenerative disorder characterized by the degeneration of dopaminergic cells in the midbrain and subsequent dopamine depletion in the striatum^{46–48}. Given the cognitive impairments reported as associated with PD, attention deficits could serve as a highlighted characteristic of the affected brain^{49–51}. Specifically⁵², used a three-stimuli oddball task to demonstrate that PD patients exhibit differences in the distractor-elicited P300 component compared to healthy controls. This difference is attributed to impairments in basal ganglia circuits, which function as modulatory circuits for fronto-parietal networks and are responsible for selecting task-relevant information and suppressing task-irrelevant information. Additionally, habituation disorders^{53,54} and habit formation disorders^{55–57} serve as other forms of cognitive impairment observed in PD patients. Furthermore, as a study investigating EEG signal properties such as the P300 and N200 components in an oddball task found, PD patients exhibit a delay in the N200 response which is associated with the slowed automatic cognitive processes reported for these patients⁵⁸.

In terms of the brain circuitry involved in learning, the basal ganglia, and specifically the striatum, controls the flow of information between cortical circuits, such as the posterior and frontal cortices, and the subcortical cortex, including the hippocampus, in support of learning^{59–61}. It also influences sensory information processing in favor of task-relevant information by modulating the fronto-parietal network, employing attention mechanisms^{62,63}. In general, the striatum functions as a dynamic controller responsible for integrating sensory information and historical information from previous outcomes to shape desired behavioral policies throughout a task^{64–69}. In mathematical terms, it is responsible for forming conditional distributions over potential actions, representing behavioral policies over time, whether they involve motor actions⁷⁰, updating working memory⁷¹, or are related to cognitive control in general^{72,73}.

Given the critical role of the striatum in dynamically shaping behavior through the learning process, disorders of attention and habituation in PD patients, as two interrelated mechanisms involved in behavioral control during learning, can be associated with impairments in striatal circuits. Moreover, considering the potential relationship between complexity and the control of uncertainty – describable in mathematical terms as an aspect of behavioral control – it follows that impairments in striatal circuits in PD patients should lead to disruptions in the entropy of conditional distributions during learning. As a result, the dynamic patterns of complexity during the sequence learning process are expected to differ in a meaningful manner between PD patients and healthy controls.

In this study, we examine the link between the complexity of brain activity and information metrics associated with sequence learning³, focusing on uncertainty and different definitions of surprise. Our findings reveal a significant superiority of the correlation between brain complexity and uncertainty in the Bayesian model compared to its correlation with other surprise metrics. Furthermore, our findings show that brain complexity follows a habituation trend during learning. Additionally, we observe a meaningful difference in the behavior of the proposed measure of complexity between PD patients and healthy controls. Based on this difference and through suggesting a quantitative assessment of the relationship between complexity and the process of habituation or attention reduction during the learning procedure, our results provide an explanation for the attention deficit or habituation disorders reported for PD patients.

Results

Brain complexity and uncertainty

To explore the relationship between complexity derived from brain activity and information-theoretical measures from the Bayesian learning premise, we evaluated the correlation between the brain's measures of complexity and various surprise metrics. Specifically, we examined the correlation of the brain activity complexity with information-theoretical metrics reflecting the puzzlement and enlightenment surprises³⁴, as well as a new measure of uncertainty in the input sequence, which along with Shannon surprise, serves as a representation of the notion of puzzlement surprise. We define this measure of uncertainty as the entropy of the posterior distribution over the stimulus space inferred through the Bayesian learning process.

Figure 1 illustrates boxplots of correlation values between brain activity-derived complexity for each of the CTL and PD groups and the puzzlement surprise measures of uncertainty and Shannon surprise. The Lempel–Ziv complexity (LZC)¹¹ and sample entropy (SampEn)¹² methods were used to measure the brain complexity. Our findings, as illustrated in Fig. 1, reveal that the correlation values for uncertainty consistently surpass those for Shannon surprise (p -value < 0.05 ; paired t -test).

Furthermore, we examined the correlation of the difference between uncertainty corresponding to each trial and that of its preceding trial with the changes in measures of brain complexity between the two trials. This was done to assess whether fluctuations in the input uncertainty could entice the brain to update its model and lead to brain activities associated with the enlightenment surprise. In a similar fashion, we also examined the KL divergence on the parameter space, which is known to serve as a representation of the enlightenment surprise³⁴, as well as implemented a new representation of this surprise based on applying the KL divergence to the stimulus space (see Methods for a description of this measure). Figure 2 illustrates boxplots of correlation values between changes before and after trials for the two measures of brain activity complexity and each of these three measures of enlightenment surprise. As indicated by these plots, the correlation values associated with difference-of-uncertainty demonstrate higher values compared to those corresponding to the KL-based measures of enlightenment surprise (p -value < 0.05 ; paired t -test).

To validate these correlations and determine which measure of uncertainty and which type of surprise in Bayesian learning exhibit significant positive correlation with brain complexity, we conducted a permutation analysis (as described in Methods) for both surprise and uncertainty, generating surrogate correlations. Then, to assess the hypothesis that each type of surprise or uncertainty had a higher correlation than its corresponding surrogate correlation, we applied a one-sided t -test. Table 1 presents the p -values for both the CTL and PD groups. The correlations for uncertainty and difference-of-uncertainty almost always yielded p -values < 0.05 ; one-sided t -test with unequal variances. In contrast, all other methods yielded non-significant p -values (> 0.05 ; using the same t -test). This suggests that only uncertainty and difference-of-uncertainty meaningfully correlate with complexity and changes of complexity between trials, respectively.

To extend our analysis to assess individual channels instead of merely focusing on average complexity across channels, and in order to derive the spatial distribution of complexity across different brain regions, we repeated all correlation analyses for complexities derived from each channel. This resulted in the topographic maps presented in Fig. 3, which illustrate the group mean correlations for each of the CTL and PD groups for both the LZC and SampEn measures of complexity, along with p -values from the t -test analysis of perm2, as discussed in Table 1. Each column of topographic maps corresponds to a distinct information metric as indicated on the figure. Additionally, for each subject group, four rows of topographic maps are presented. The first and third rows in each subject group set illustrate topographic maps corresponding to the group mean values of SampEn and LZC, respectively. In each row, the topographies are separated into two sets based on their information metrics being a representative of the puzzlement or enlightenment notions of surprise. As observed in the first and third rows for each subject group, uncertainty, as a measure of puzzlement surprise, consistently shows more significant correlations across most channels compared to Shannon surprise. Similarly, within the representations of enlightenment surprise, difference-of-uncertainty consistently demonstrates a higher correlation with brain complexity than the KL-based metrics. To assess the significance of these correlations in a graphic fashion, the second and fourth rows for each subject group correspond to the logarithm of the p -values from the t -test. The regions highlighted in bright to dark red shades indicate channels showing significant correlations based on permutation analysis. Notably, for both subject groups and data blocks, significant correlations with brain complexity were observed only for uncertainty and difference-of-uncertainty. As these plots indicate, the patterns of channels with significant correlations are predominantly distributed over the frontal, central, and parietal regions.

Complexity trends and habituation

To examine the dynamical pattern of brain activity complexity during sequence learning, we calculated the complexity of brain activity after the onset of each stimulus in the sequence by two methods: Lempel–Ziv complexity (LZC)¹¹ and sample entropy (SampEn)¹². We then plotted the grand average of the derived time

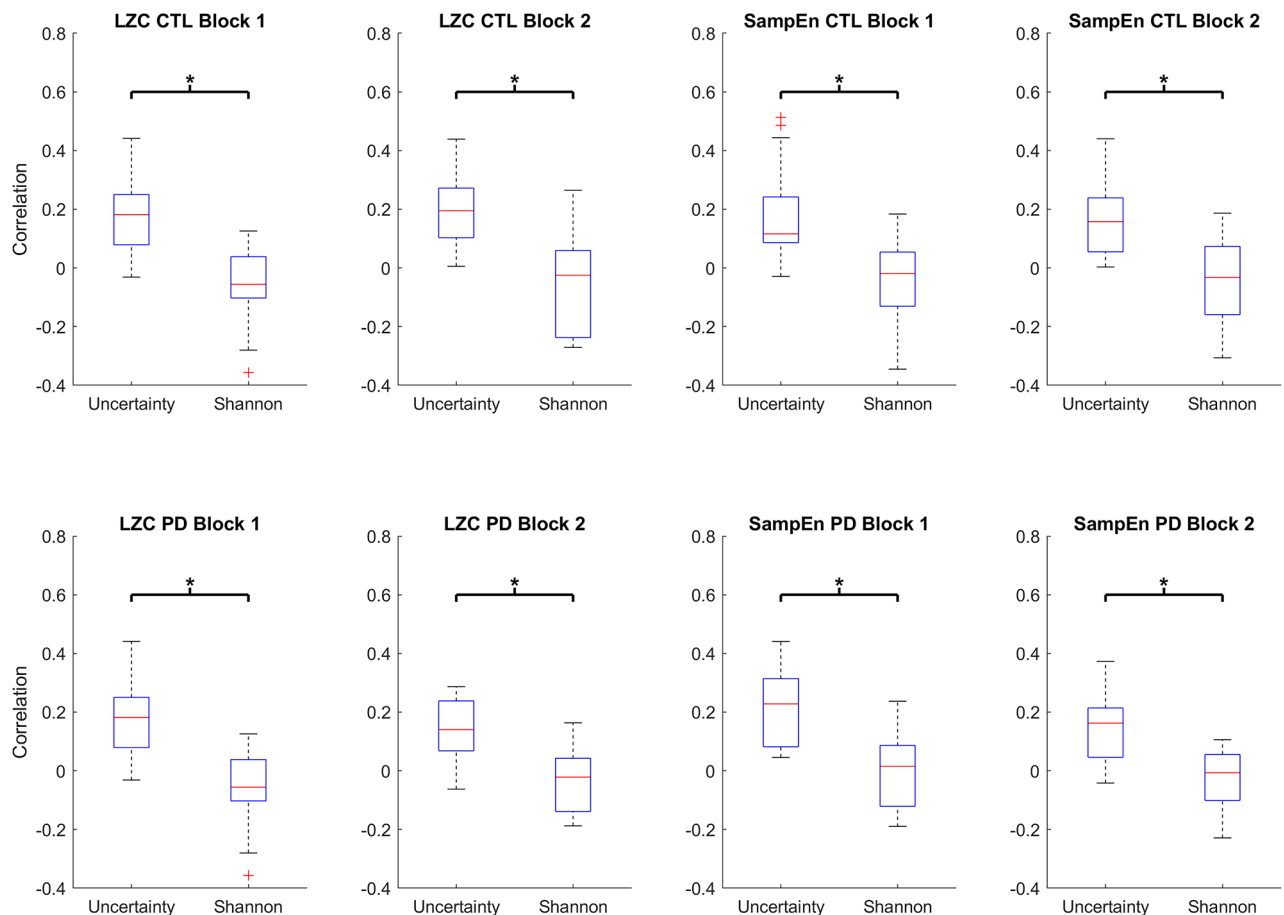


Fig. 1. Boxplots of correlation of brain activity-derived complexity with uncertainty and Shannon surprise, both of which are surprise metrics calculated from Bayesian inference. In each trial of the oddball task, two measures of brain complexity, Lempel–Ziv complexity (LZC) and sample entropy (SampEn), are applied to each channel of the EEG time series within the 200–400 ms post-onset time window. Then, the average complexity is calculated over all channels, resulting in an array of complexity measures for each subject with the same size as the number of trials in the oddball task. On the other hand, for each information metric, a value is assigned to each trial of the oddball task based on the Bayesian learning model. The correlation between the brain complexity and each information metric is then calculated for each subject. This procedure is performed separately for each of the two blocks of the oddball task. The boxplots indicate that the correlation of complexity with uncertainty consistently surpasses that with Shannon surprise (p -value < 0.05 , paired t -test). The boxplots in the first and second rows correspond to the CTL and PD groups, respectively. The correlations shown in the boxplots, for both uncertainty and Shannon surprise, are optimized over the parameter w . This parameter models memory length for each subject. The boxplots indicate the median (red line inside the box), the interquartile range (box edges for Q1 and Q3), and whiskers extending to the minimum and maximum values. The red plus signs shown on some boxplots correspond to subjects with considerably different correlation values from their group.

series of complexity throughout the sequence of stimuli for each of the two blocks of the oddball task, as shown in Fig. 4. In the first block, the CTL group displays a decreasing trend in complexity, eventually stabilizing at a fixed value. This phenomenon is consistent with the first property of habituation which predicts a decreasing trend in brain activities to an asymptotic level during learning⁴⁰. The PD group also shows a similar decreasing trend. However, the stabilization of complexity for the PD group occurs later in the stimuli sequence compared to the CTL group. This difference in trends manifests itself as an increasing divergence of the PD group's complexity from that of the CTL group, as highlighted by the complexity trends in Fig. 4. As can be seen in Fig. 4, the complexity trend of the PD group diverges from that of the CTL group, beginning from the middle stage and leading to a statistically significant difference between the two groups in the late stage of learning.

To highlight the decreasing behavior and convergence in complexity, we examined the trend for the slope of the line fitted to the complexity windowed over its time series. In the first block, as illustrated in Fig. 5, the initial slope of the complexity trend for both the CTL and PD groups is negative. This reflects a decreasing trend in complexity. For the CTL group, the slope approaches zero around the midpoint of the stimulus sequence and stays near zero until the end of the sequence. In contrast, for the PD group, the slope only reaches zero towards

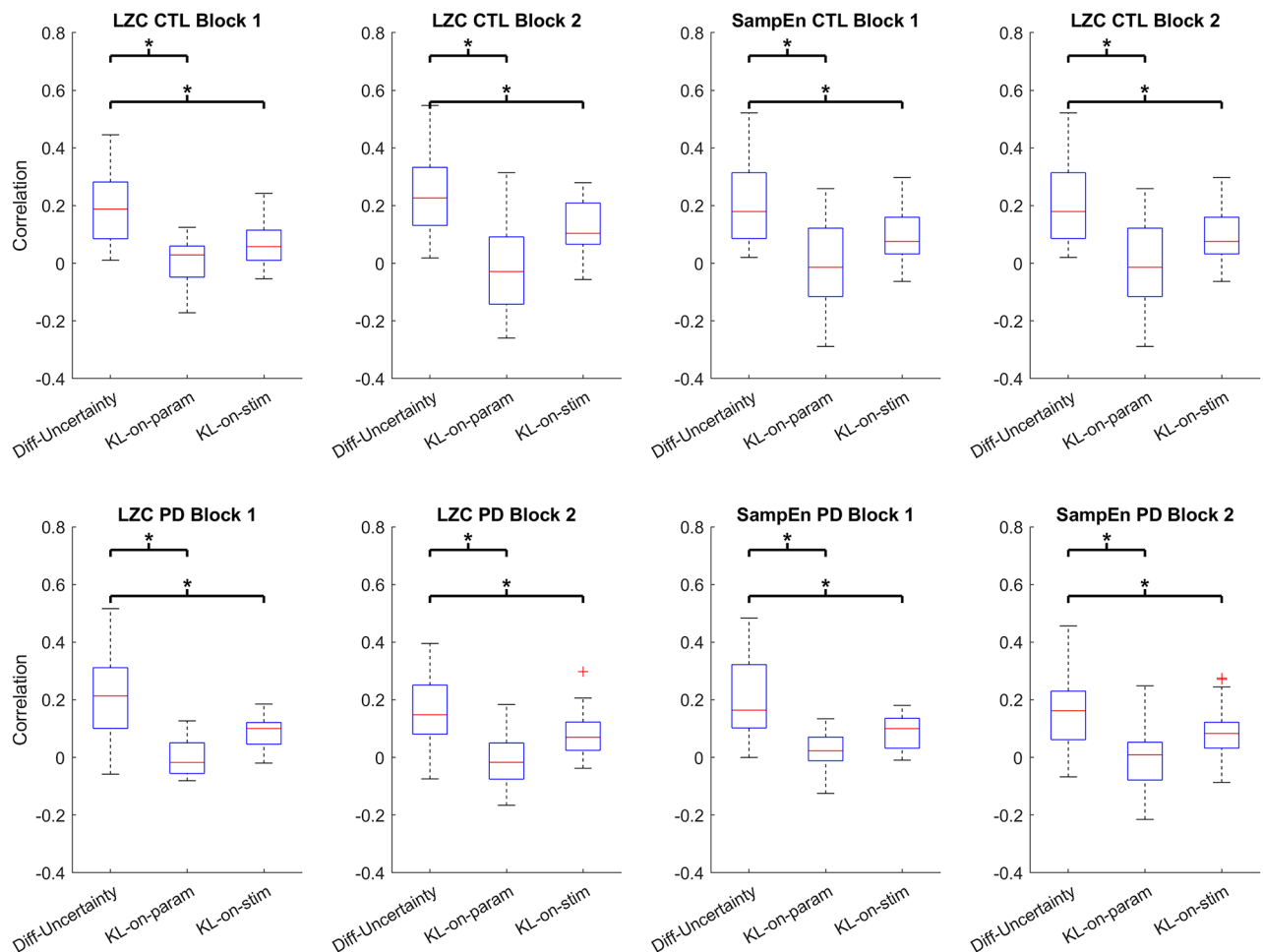


Fig. 2. Boxplots of correlation of changes in brain activity-derived complexity before and after stimulus with difference-of-uncertainty, KL divergence on parameter space, and KL divergence on stimulus space, all associated with the discrepancy between the prior and posterior distributions before and after observing a stimulus in the sequence. These distributions are either defined over the parameter space (as in KL on parameters) or the stimulus space (as in KL on stimulus and difference-of-uncertainty). The boxplots illustrate the correlation of these measures of surprise with the change in brain complexity from one trial to the next (calculated as the *next* trial minus the *current* trial). The boxplots in the first and second rows correspond to the CTL and PD groups, respectively. Notably, the correlations associated with difference-of-uncertainty exceed those for KL on stimulus and parameter spaces (p -value < 0.05 ; paired t -test). Correlations for all surprise types are optimized using the parameter w , which represents memory length for each subject. The boxplots indicate the median (red line inside the box), the interquartile range (box edges for Q1 and Q3), and whiskers extending to the minimum and maximum values. The red plus signs shown on some boxplots correspond to subjects with considerably different correlation values from their group.

the sequence's final stages, suggesting a later convergence of complexity for the PD group in the first block. In the second block, however, the trends in the slope for the CTL and PD groups do not show significant differences, as there is no statistical difference between the slope of complexity trends throughout the stages of learning.

To extend the analysis applied to the trend and slope of complexity in order to assess the spatial behavior of complexity, we grouped the channels into four sets corresponding to the frontal, parietal, central, and temporal regions. The complexity for each trial was averaged over the channels within each of these regions, resulting in four regional complexity trends. Subsequently, by performing the same windowing and line fitting analyses, we examined the trend and slope of complexity for both the LZC and SampEn complexity measures for each region. As illustrated in Fig. 6, the trend of complexity for the CTL group across all regions and for both measures of complexity exhibits convergent behavior. Additionally, the slope of complexity for the CTL group consistently starts with a negative value and approaches zero after the beginning stage of learning and remains around zero until the final stage of learning in each block. Moreover, the same pattern of difference observed in Figs. 4 and 5 also exists between the CTL and PD groups in the regional complexity trends. In the first block, the convergence of complexity for the PD group occurs at a later stage for all regions. Furthermore, the slope of complexity for the PD group in the first block remains lower than zero for more trials compared to the CTL group. However,

Data Block	Surprise and uncertainty measures	Perm1 LZC	Perm2 LZC	Perm1 SampEn	Perm2 SampEn
CTL group					
1	Uncertainty	6.98×10^{-4}	0.0028	0.0064	0.018
	Difference-of-uncertainty	0.0012	7.10×10^{-5}	0.0026	3.57×10^{-4}
	Shannon surprise	0.99	0.99	0.99	0.99
	KL on stimulus	0.99	0.76	0.99	0.92
	KL on parameters	0.99	0.99	0.98	0.99
2	Uncertainty	2.18×10^{-4}	7.26×10^{-5}	0.0062	0.017
	Difference-of-uncertainty	2.42×10^{-4}	4.70×10^{-5}	0.0046	0.0012
	Shannon surprise	0.99	0.99	0.99	0.99
	KL on stimulus	0.22	0.08	0.82	0.39
	KL on parameters	0.98	0.97	0.99	0.99
PD group					
1	Uncertainty	7.79×10^{-5}	0.0016	4.57×10^{-5}	0.0013
	Difference-of-uncertainty	0.0033	4.36×10^{-4}	0.0014	1.24×10^{-4}
	Shannon surprise	0.99	0.99	0.99	0.99
	KL on stimulus	0.89	0.47	0.81	0.18
	KL on Parameters	0.99	0.99	0.99	0.99
2	Uncertainty	0.0097	0.0048	0.022	0.0099
	Difference-of-uncertainty	0.027	4.40×10^{-6}	0.093	0.024
	Shannon surprise	0.99	0.99	0.99	0.99
	KL on stimulus	0.95	0.66	0.89	0.47
	KL on Parameters	0.97	0.99	0.98	0.99

Table 1. P-value table for hypothesis testing to assess the significance of correlation between surprise metrics from Bayesian inference and brain activity-derived complexity. In the Perm1 test, the complexity sequence for each subject undergoes 200 random permutations. The correlation is then calculated and optimized over the parameter w for each permutation. In the Perm2 test, 200 random permutations are applied to the stimulus sequence for each subject. Subsequently, the surprise time series are recalculated, and the correlations are again determined and optimized over the parameter w . For each test, the correlations resulting from the permutations, totaling 5,000 (200 permutations $\times$ 25 subjects), are compared to the original correlation set, which consists of 25 correlations, using one-tailed t-test with unequal variances. As indicated in bold font, only uncertainty and difference-of-uncertainty (calculated as *after* minus *before* the stimulus) exhibit significant correlations with p-values less than 0.05.

this delayed convergence in the trend of complexity for the PD group is not observed in the second block, and no significant difference exists with the CTL group. Moreover, based on the p-values resulting from t-test analysis of the slope of complexity, it can be noted that the frontal region exhibits more disparity between the CTL and PD groups in the trend of complexity, showing p-values lower than 0.001 marked by three stars in Fig. 6a.

Spontaneous recovery

The second property of habituation discussed in⁴⁰ is spontaneous recovery. This property refers to the partial recovery or resurgence of a brain response characteristic after a pause in the presentation of the sequence of stimuli for a period of time. In the oddball task setup, a break occurs between the first and second blocks, during which the presentation of stimuli is halted. To investigate the effect of this property on the complexity of brain activities, we compared the average of the last ten complexity values in the first block to the average of the first ten complexity values at the start of the second block. To this end, we evaluated the hypothesis of spontaneous recovery, which stipulates a higher level of complexity occurs at the beginning of the second data block compared to the end of the first block, using paired t-test. As shown in Fig. 7, our results reveal that the complexity values, for both LZC and SampEn, are higher in the beginning of the second block compared to those at the end of the first block for both the CTL group (two-sided paired t-test: LZC: p-value = 0.022, $t(24) = 2.449$, SampEn: p-value = 0.069, $t(24) = 1.904$) and the PD group (two-sided paired t-test: LZC: p-value = 0.001, $t(24) = 3.74$, SampEn: p-value = 0.002, $t(24) = 3.467$). The CTL group's SampEn complexity comparison was marginally non-significant ($p > 0.05$). These findings suggest the occurrence of a recovery in complexity values after pausing the presentation of the stimuli sequence.

Discussion

Recent studies have underscored the significance of brain complexity in shaping cognitive abilities^{2,4,5}. The relationship between complexity and the structure of information flow in the brain was studied in⁴ through

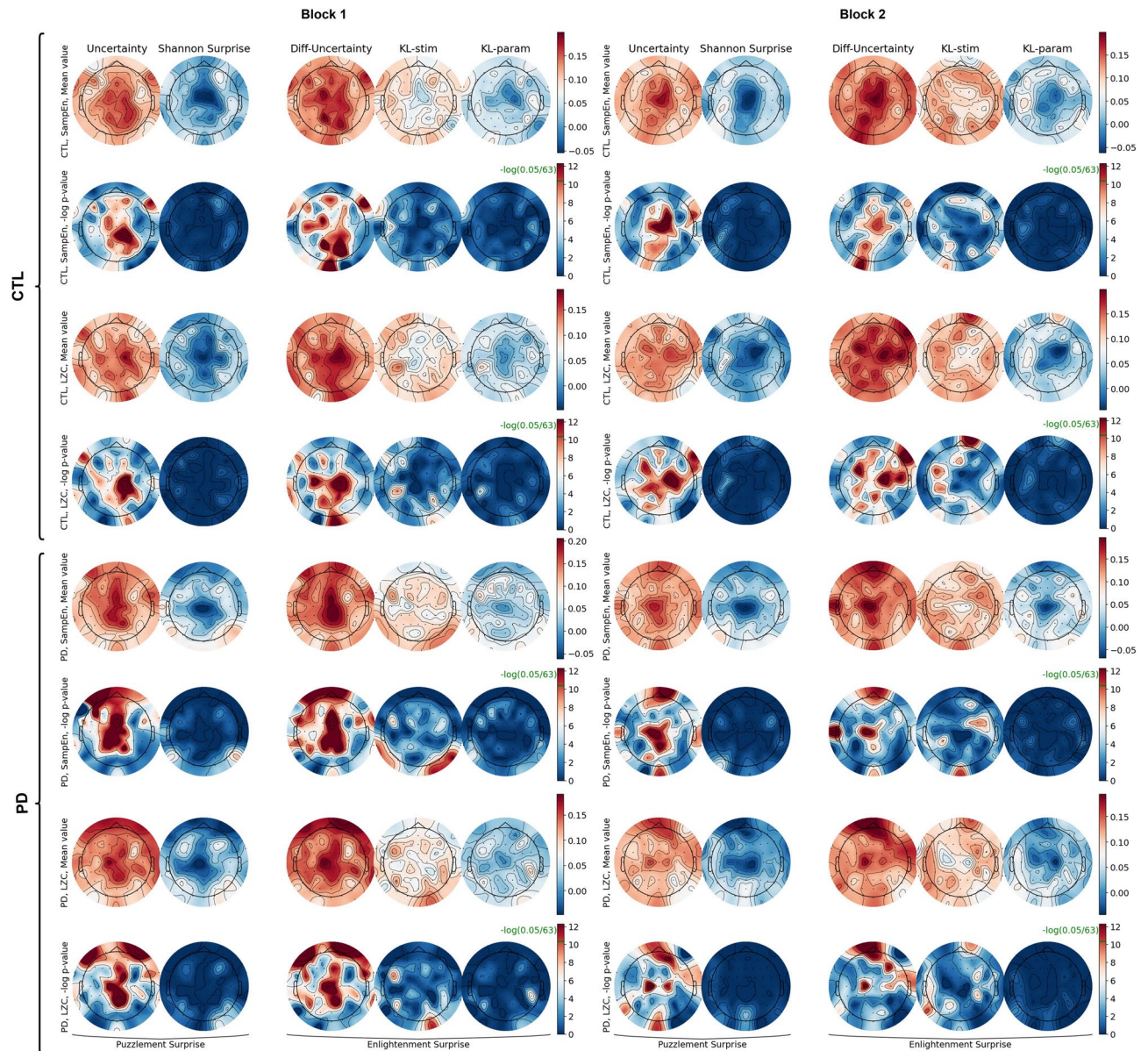


Fig. 3. Topographic maps of correlations between surprise metrics from Bayesian inference and the complexity of brain activity derived from each channel and topographic maps of the corresponding p-values. For each subject group and oddball block, five columns of topographic maps are presented, each corresponding to one information metric. These topographies are categorized into two groups that represent the puzzlement or enlightenment notions of surprise. For each subject group, the first and second rows respectively contain topographies of the group mean correlation for the SampEn complexity and the corresponding p-values ($-\log(p\text{-value})$) resulting from the perm2 test over channels. The third and fourth rows include topographies corresponding to the group mean correlation of the LZC complexity and the corresponding p-values ($-\log(p\text{-value})$), also resulting from the perm2 test over channels. It can be observed that, in the group of puzzlement surprises, for both subject groups, task blocks, and complexity types, the group mean correlation for uncertainty—as a measure of Bayesian surprise—is significantly higher than that for Shannon surprise over most brain regions, and especially in the frontal, central and parietal regions. Moreover, in the group of enlightenment surprises, the group mean correlation for difference-of-uncertainty is noticeably higher than that for other metrics (KL-stim and KL-param) across both subject groups, task blocks, and complexity types. Additionally, it should be noted that, only for uncertainty and difference-of-uncertainty, there are regions with higher levels of correlation than the significance level, which are mainly distributed over the frontal, central and parietal regions. The significance level is indicated by a green line plotted on the color bar of the $-\log(p\text{-value})$ maps, which corresponds to the significance level $-\log(0.05/63)$ after correction, where 63 is the number of channels. Colors higher than this mark indicate regions with a correlation higher than the significance level.

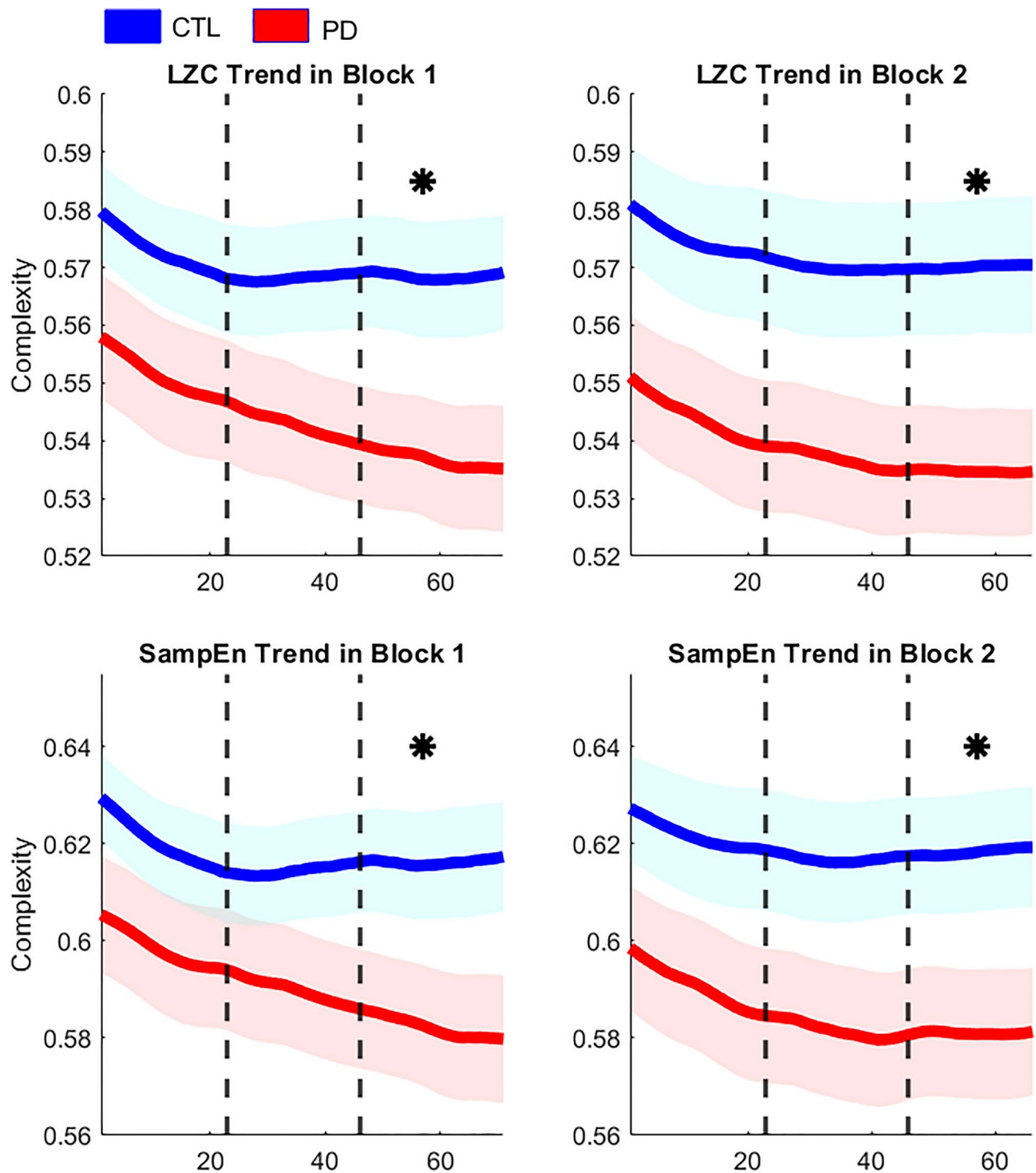


Fig. 4. Trends of complexity measures over the two blocks of the input sequence. The LZC and SampEn complexities were calculated for each subject in each trial. The complexity values underwent filtering using a moving average window of size 30 to eliminate rapid fluctuations. For both groups, the means are depicted in the first and third rows along with standard error of the mean (SEM) (red represents the PD group and blue denotes the CTL group). To compare the trend of complexity through learning between the PD and CTL groups, the time bins were divided into three stages of learning: early (time bins 1:23), middle (time bins 24:46), and late (time bins 47:70). At each stage, two-sided t-test was applied to compare the average complexity of the PD and CTL groups over the trials of that stage. Each stage showing a p-value lower than 0.05 or 0.01 is marked with one or two stars, respectively. As can be seen, for all blocks and complexity measure, the trend of complexity of the PD group significantly diverges from that of the CTL group in the late stage of learning.

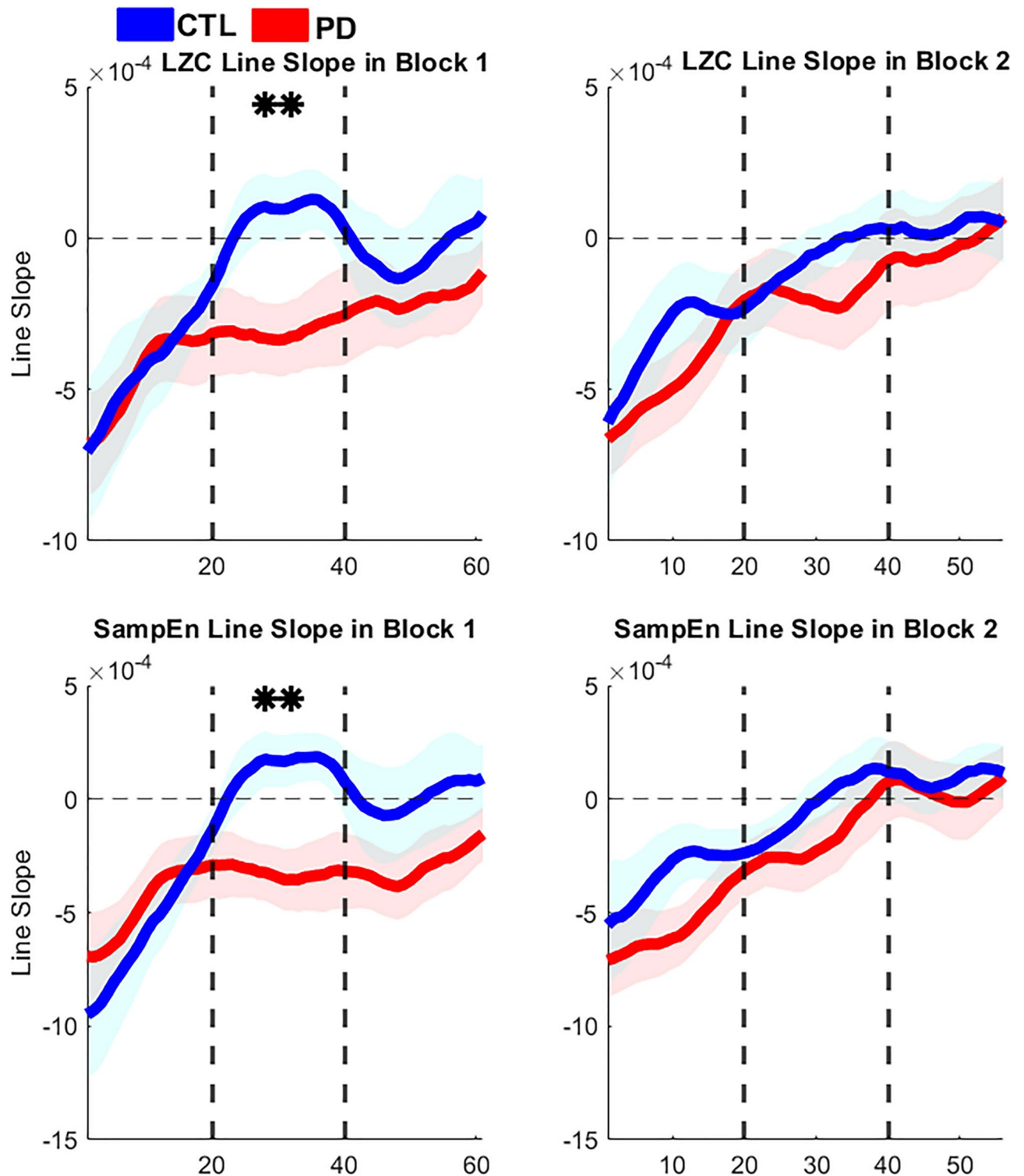


Fig. 5. Trends of the slope of complexity measures over the two blocks of the input sequence. For each subject in each trial, the LZC and SampEn complexity measures were calculated. The trials were then segmented into windows of size 40, and a line was fitted to each window. The means and their SEM shades are displayed in the first and third rows (red represents the PD group and blue denotes the CTL group). To compare the trend of complexity slope between the PD and CTL groups through the learning process, the time bins were divided into three stages of learning: early (time bins 1:20), middle (time bins 21:40), and late (time bins 41:60). Each stage showing a p-value lower than 0.05 or 0.01 is marked with one or two stars, respectively. As can be seen, in the first block, the slope of the complexity trend for the CTL group diverges from that of the PD group in the middle stage, approaching zero and remaining there till the end of learning. Nevertheless, the slope of complexity trend for the PD group remains negative until the late stage where the slope approaches zero.

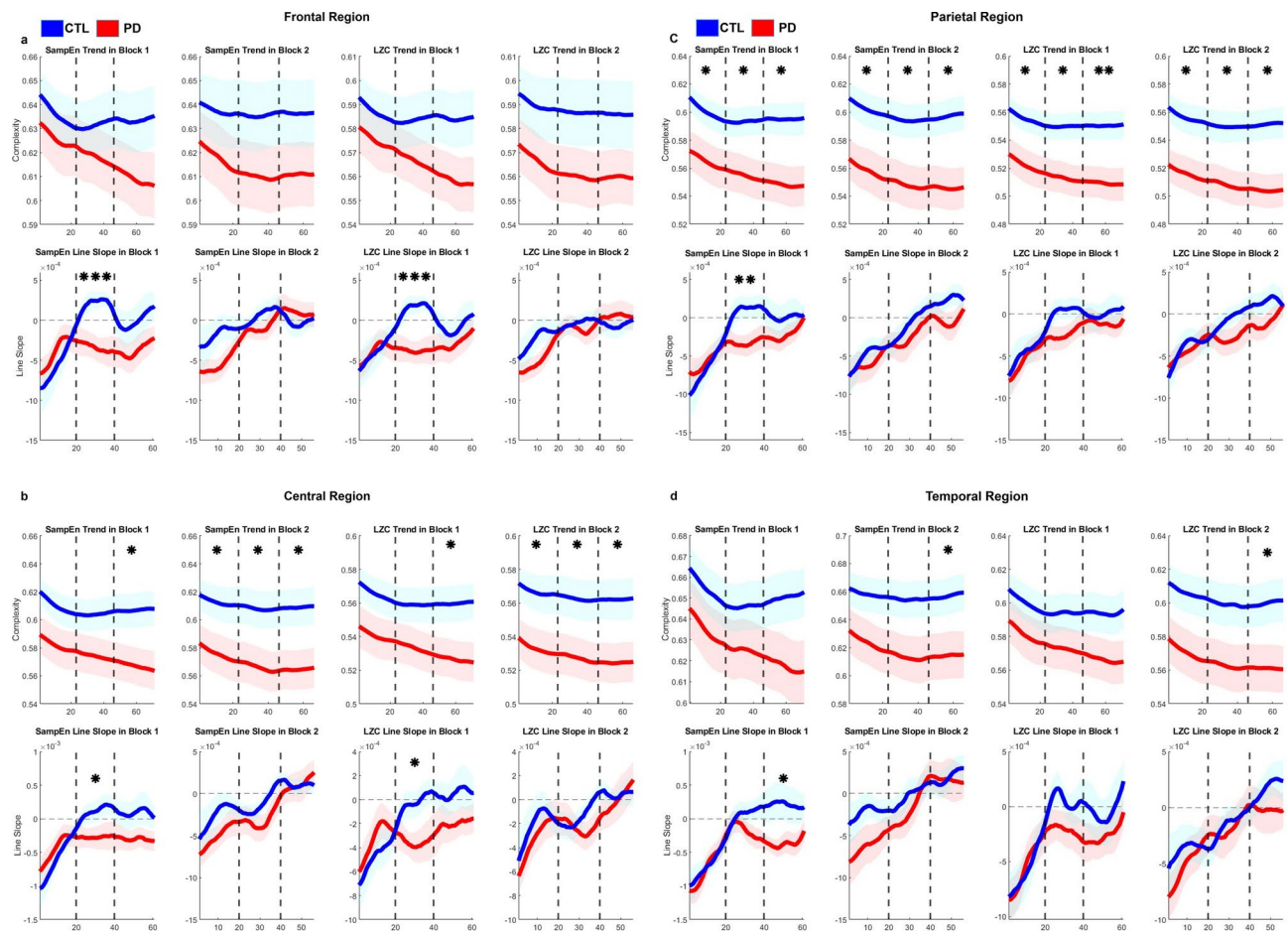


Fig. 6. The trends of complexity and their slopes for a) frontal, b) central, c) parietal, and d) temporal regions. For each region, channels are selected based on the electrode naming in the 10–20 standard. For example, all channels with names starting with ‘F’ are selected as channels representing the frontal region. The complexities associated with each channel for each trial are averaged over each regional cluster of channels. Then, the trends of the mean complexity through trials are plotted for each region for the PD and CTL groups. Additionally, following the same procedure discussed for Figs. 4 and 5, the trends of the slope of complexity for each region as well as their SEMs are plotted. The same pattern of convergence in the complexity for the CTL group as the overall complexity shown in Figs. 4 and 5 is also seen in each brain region. Furthermore, the late convergence of the complexity trend for the PD group in the first block is also manifested in the trend of complexity slope in all regions.

analyzing connectivity and temporal dynamics. In addition, information flowing due to the dynamics of brain activities across various regions was shown to play a pivotal role in establishing cognitive control and adaptive behavior during cognitive tasks such as learning^{7,74,75}. A recent study even used the complexity of EEG data to develop a classifier for EEG signals according to varying levels of attention during sustained attention tasks⁷⁶. A large body of research on brain activity complexity also delves into its relationship with the state of consciousness^{25,43}. Despite the connection established between complexity and information processing in the brain^{2,4}, there remains a lack of knowledge about how the complexity of brain activities contributes to the formation of different aspects of cognitive functions, particularly the learning process, habituation, and habit formation.

In the context of learning, complexity is often considered as an emergent property arising from structural interactions between brain regions within a network. This network undergoes modifications during the learning process or behavioral adaptation, caused by changes in synaptic strengths. For example, functional connectivity metrics were considered in⁷⁷ as measures of spatial complexity to identify dynamic changes in the modular structure of the brain during the learning process. Another study emphasized the role of synapses as the physical substrate for the interaction between brain regions in forming brain complexity and facilitating advanced computation⁷⁸.

Despite these findings, an exact, intuitive relationship between the complexity of brain activities and the dynamic, high-level aspects of the learning process remains elusive. This includes the uncertainty associated with the conditional distribution at each stage of the learning process, and its convergence to a desired conditional distribution. In this study, we explored the dynamic pattern of the complexity of brain activities in terms of the

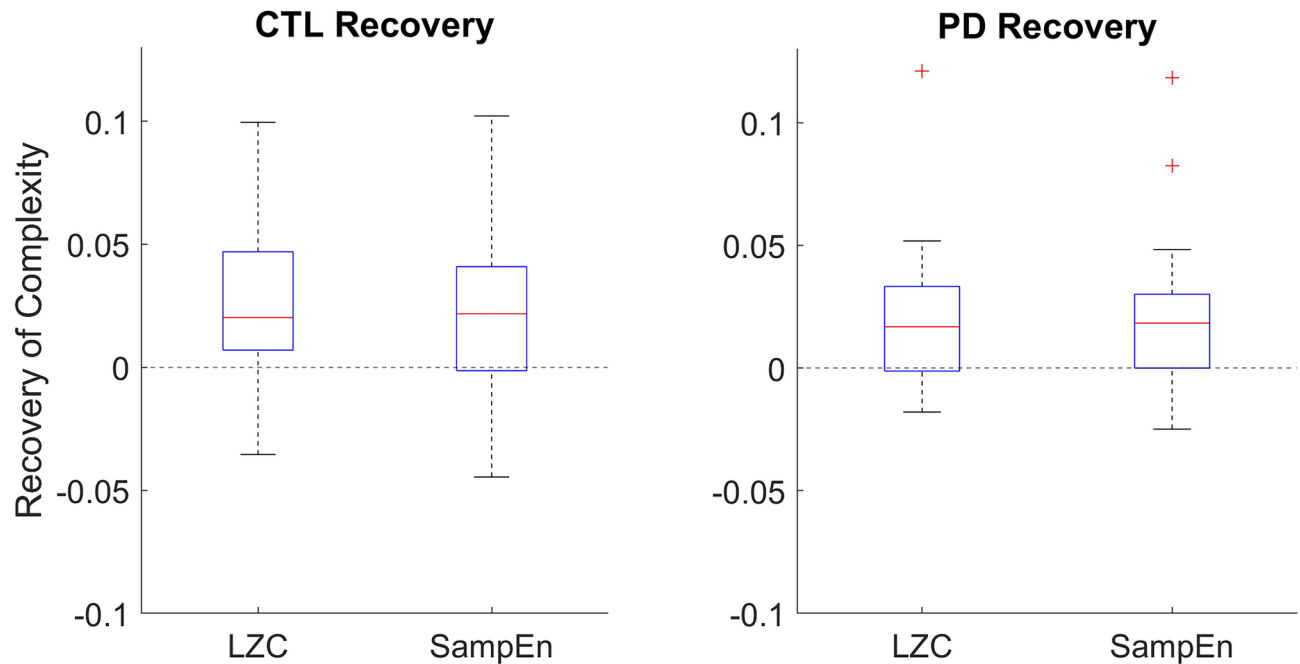


Fig. 7. Boxplots illustrating differences between the averages of the initial 10 complexity values in the second block and the final 10 complexity values in the first block. As can be seen, complexities at the start of the second block are consistently greater than those at the end of the first block for both PD and CTL groups and for both complexity measures. The paired two-sided t-test p-values for the CTL group are 0.069 and 0.022 for SampEn and LZC, respectively. Additionally, the paired two-sided t-test p-values for the PD group are 0.002 and 0.001 for SampEn and LZC, respectively. The red plus signs shown on some boxplots correspond to subjects with considerably different correlation values from their group.

rate of information generation during brain dynamics within a sequence learning process. We also examined the dynamic information-related measures involved in the mathematical models of sequence learning³, such as various types of surprise as well as uncertainty serving as the entropy of the conditional distribution formed during sequence learning for each stimulus given the previous stimuli. Through this analysis, we observed a noticeable correlation between uncertainty in Bayesian learning and brain complexity throughout the learning process. Furthermore, we identified a delayed convergence in the brain complexity trend for the PD group, as shown in Figs. 4, 5, and 6. Given this meaningful difference between the PD and CTL groups, as well as the relationship between complexity and uncertainty in the Bayesian learning process, we propose that cognitive disorders in PD such as habituation disorder^{53,54}, habit formation disorder^{55–57}, and slow cognitive processing^{79,80} may be attributed to impaired formation of the conditional distribution, which manifests as the deficiency of the brain in uncertainty reduction throughout these cognitive processes.

Brain complexity is represented by uncertainty and not novel information in Shannon or Bayesian models of input surprise

The learning process primarily involves handling and processing input information to form a desired behavior within a task to be learned. In mathematical terms, information processes such as encoding or filtering the information could potentially be associated with shaping a distribution of potential actions. These actions could include a set of possible decisions which can be made in an environment, potential outcomes which should be predicted, or even potential models for describing input data, all based on the input information model derived from previous observations. Furthermore, perception can be described as calculating the posterior distribution of the potential causes of sensory inputs, and perceptual learning as updating the prior distribution of these causes⁸¹. Intuitively speaking, learning is driven by the existence of some degree of uncertainty about the optimal behavior which needs to be reduced during the learning process. Therefore, gathering and encoding information during the learning process can result in a reduction of uncertainty. Moreover, uncertainty arises from the degree of randomness, which in turn originates from a lack of information from the learner or observer's perspective. Without any randomness, both uncertainty and subsequently gathering information may lose their meanings. Mathematically speaking, uncertainty about the desired behaviors can be characterized by properties associated with the distribution of potential behaviors and how this distribution is shaped during the learning process. In the context of sequence learning, learning involves understanding the mechanism by which the input sequence is generated, or being able to predict the next stimulus using the stochastic model for sequence generation learned from previous stimuli. This again turns out to form and shape the distribution of potential parameters determining the stochastic model of the input or the potential states for the next item in the sequence.

Many mathematical models for sequence learning assume the Markov property on the generation of stimuli in the sequence^{3,30}. Based on this assumption and using a Bayesian learning approach³, suggests that the brain as a near-ideal observer attempts to estimate a distribution for the transition probability of the input stimuli. In this mathematical model for learning, there are several information-related measures which model high-level aspects of learning such as the surprise level associated with each stimulus or the uncertainty about the next stimulus, both of which are calculated from the distribution formed during the presentation of the sequence. Based on the definition of surprise for a stimulus as its unlikelihood, Shannon surprise is calculated as the negative logarithm of the probability associated with the stimulus in the internal model or the learned distribution of all potential stimuli⁸². This measure quantifies the level of conflict between the internal model and the world, and is termed as puzzlement surprise³⁴. In an informational view, surprise can also stand for the amount of new information encoded by the learner due to observing a specific stimulus. From this perspective, surprise can be defined by the amount of change which the internal model undergoes after observing the current stimulus, and is termed as enlightenment surprise³⁴. Given the informational view, a model for the enlightenment surprise is introduced in^{83,84} and termed as Bayesian surprise. Bayesian surprise is formulated as the Kullback–Leibler (KL) divergence between the prior distribution of the model parameters and the posterior distribution calculated after observing the current stimulus. The idea underlying Bayesian surprise can be extended to define a new quantification of surprise through capturing the amount of change in the internal model by calculating the KL divergence between the prior distribution formed for the stimulus space, not the model parameter space, and the corresponding posterior distribution. This new quantification of surprise may potentially be more biologically tractable as it has to do with a distribution for the stimulus space which consists of a finite set of discrete items, i.e. the stimuli types, rather than a distribution for a continuous variable which models the transition probability. Furthermore, this measure of surprise is pertinent to the prediction of a subject about the next stimulus, making it less dependent on a particular computational learning model consisting of assumed model parameters.

Along with other information-theoretical measures calculated from the conditional distribution formed during the learning process, we also calculated the entropy of the conditional distribution of the stimuli space formed after observing each stimulus to measure the remaining uncertainty of the learner about the stimuli after each observation. Entropy captures the optimal average code length for describing outcomes of a process given a specific distribution for generating them. Given this view, and intuitively speaking, the entropy of the conditional distribution formed by a learner after observing some stimuli is an indicator of the average code length prescribed by its current coding scheme driven by the learned internal model. Initially in the sequence learning process, the learner's coding scheme for encoding the stimuli produces a higher code length than the optimal length, reflecting higher uncertainty about the stimuli. The more stimuli the learner observes, however, the closer the internal model becomes to the true model of sequence generation and this results in lower uncertainty. Therefore, as the internal model gets closer to the underlying sequence generation model, the coding scheme produces a lower average code length which is closer to the optimal code length and reflects the lower uncertainty about the stimuli. Refs^{85,86} discuss the role of uncertainty in neural coding schemes by considering uncertainty as a property associated with the conditional distributions obtained from Bayesian approaches. Developing a Bayesian model of neural representation, Friston²⁷ considers the entropy of the sensory state (stimuli) distribution as average uncertainty about the stimuli generation process. One plausible explanation justifying all these perspectives is that the more uncertain the brain is about an item in the input sequence, the more new information it requires to be encoded during the observation of the sequence to reduce this uncertainty, resulting in a coding scheme with a higher average code length. From this viewpoint, we posit that the reduction in uncertainty between the prior and posterior distributions, respectively associated with before and after observing a stimulus in a sequence, can be considered as an enlightenment surprise³⁴.

In addition to information-theoretical measures in the Bayesian learning model which are applied to the input sequence distributions to describe the amount of uncertainty or randomness, other quantification measures such as Lempel–Ziv complexity (LZC)¹¹ and sample entropy (SampEn)¹² can be employed to capture the level of information processing in the brain as observing the sequence of input stimuli results in a form of activity dynamics in time. These methods attempt to quantify the rate of new patterns generated in brain activity data recorded in the form of time series. Intuitively speaking, these new patterns can be considered to reflect the rate of new information from the environment encoded by the brain. In other words, the new patterns in brain activity could potentially be linked to the coding scheme used by the brain to encode the new information.

Given this outlook, higher levels of new information, regarded in terms of the code length from a Bayesian learner's perspective, could be related to higher numbers of new patterns generated in the brain activities as a Bayesian learner⁸⁶ or an ideal observer⁸⁷. To assess this hypothesis, we investigated the correlation between each of the LZC and SampEn measures of complexity derived from the brain data and the information-related measure of uncertainty, i.e. the entropy of the conditional distribution formed for the stimulus space by the Bayesian learner²⁷. Figures 1 and 2 highlight the stronger association of uncertainty with complexity measures (LZC and SampEn) compared to Shannon surprise and KL divergence. Moreover, the superiority of both uncertainty and difference-of-uncertainty across the scalp can be seen in topographic maps in Fig. 3.

To explain the superiority of entropy in describing the complexity of brain data, it is important to note that the measure used for entropy, which captures the level of uncertainty, incorporates information from all stimuli, rather than just the recently-observed one. In contrast, both Shannon surprise and Bayesian surprise (calculated as KL divergence in our analysis) respectively measure the amount of new information or the necessary information in the recently-observed stimulus for updating the internal model. Therefore, measures based on entropy are expected to demonstrate higher consistency with the notion of complexity as an indicator of the level of information integration². Furthermore, only entropy captures the average code length of the coding scheme for all stimuli space developed by a Bayesian learner. Therefore, entropy could potentially better reflect the level of new pattern generation due to the brain's dynamics, consistent with the hypothesis mentioned earlier.

In addition, the selected time window after the stimulus onset, during which the rate of new pattern generation in EEG data is calculated through LZC and SampEn measures, relates to higher-order cognitive processes which involve historical information and not just the information related to the current stimulus³⁰. Argues that brain responses later in the epoch better reflect the formation of transition probability inferences than the early responses which are more responsive to simpler statistics such as the item frequency in an oddball experiment. Furthermore⁸⁸, posits that late brain responses are more sensitive to global regularity in the sequence of auditory stimuli and capture complex pattern in the sequence. Therefore, entropy, which measures the average code length for all stimuli space, can better capture the rate of new pattern generation in EEG data during this time window than the Shannon⁸² and Bayesian surprises^{83,84}, which are both more sensitive to the information related to the current stimulus in the sequence.

Brain complexity follows habituation trend during learning

As a learner progresses through a sequence, they become more familiar with its patterns, reducing the need to encode new information or update their internal model. If the brain's generation of new patterns reflects this encoding, it is expected that the rate of new pattern generation should decrease through sequence learning. Furthermore, this decrease is also expected given the relationship between complexity and uncertainty, and by noting that a learning process usually manifests as a reduction in uncertainty²⁷. In the context of sequence learning, uncertainty can be related to the unpredictability of the next stimuli, and is mathematically modeled as the entropy of the conditional distribution formed for the next stimulus given the previous ones.

Habituation, known as a manifest of the learning process^{40,41}, is defined as a decrease in response to the repetition of stimuli in a sequence. The decrease can be distinguished from sensory adaptation by its dependence on the statistical properties of the stimuli⁶⁰. Furthermore, a more complex form of learning is habit formation which is a process initiated as a goal-directed behavior that becomes automatic as a stimulus-driven behavior^{64,89,90}. In computational terms, a habit can be considered as a formed policy in favor of a task or goal, through shaping and utilizing a conditional distribution for potential actions given an observation. Given the mathematical view of learning discussed earlier, these forms of learning can be potentially linked to situations in which there is a form of convergence in the conditional distribution shaped through the learning process. Such convergence view to habituation is also consistent with the first property associated with habituation⁴⁰, pointing to a decreasing trend in brain activities in response to the sequence of stimuli towards an asymptotic level in exponential or linear decay format. In informational terms, this decreasing trend in brain activities can be potentially seen as an implication of a reduction in the rate of new information which should be encoded as a result of observing a new stimulus in the sequence. In Bayesian updating terms, this decreasing trend is equivalent to a reduction in the changes that the conditional distribution undergoes due to the observation of a new stimulus during the sequence learning process. In the habit formation process^{64,89,90}, this reduction can be seen as the convergence of the conditional distribution of actions given the observations, and be interpreted as learning a habit.

Juxtaposing the elaborated perspectives on habituation and habit formation and the relationship between the complexity of brain activities and the entropy of the formed conditional distribution for the next stimulus during sequence learning, we expect to see a form of convergence in the trend of complexity during sequence learning. Consistent with this expectation, as can be seen in Fig. 4 (trend of complexity) and Fig. 5 (slope of complexity), a decreasing trend is observed in the grand mean of complexity during the presentation of the sequence of stimuli for the CTL group, which exhibits a convergent behavior. As more support for this hypothesis, similar patterns of convergent behavior can be observed for the CTL group across different brain regions in Fig. 6. Furthermore, another property associated with habituation is spontaneous recovery^{40,41}, which argues for a recovery in the strength of neural responses after habituation when the stimulation is paused and restarted. Consistent with this property, a form of recovery can be observed by comparing the complexity trends between the end of the first block of the oddball task and the start of the second block, in which the same sequence is presented after a short period of rest, see Fig. 7.

From a biological view, basal ganglia, and especially the striatum, as part of the dopaminergic system of the brain, play a critical role in habitual control⁵⁷ and habit formation⁶⁴. In addition, basal ganglia circuits play a role in probabilistic sequence learning^{91,92}. Furthermore, the activities of the striatum in the time window between a reinforcement learning task's current outcome and the next outcome contribute to the integration of information from previous outcomes to predict the next outcome⁹³. The operation of all these circuits are impacted by the dopamine deficit due to PD. Therefore, it is expected to see changes in the complexity of brain activities associated with information integration when habits are formed in PD patients. As can be seen in Figs. 4 and 5, convergence of the complexity trend occurs with delay for PD patients in the first data block compared to the CTL group. Similar delayed convergence behavior can also be observed across different brain regions, and particularly in the frontal region, as illustrated in Fig. 6. This delayed convergence can serve as evidence for the slower performance of PD patients in automatic cognitive processes reported in⁷⁹, and for the delay in response initiation due to the delay in cognitive processes (and not motor processes) reported in⁸⁰. The severity of this disparity between the two groups in the frontal region can be potentially attributed to the critical role of the dopaminergic system in habit control, which acts through the projection of the striatal circuits onto the frontal region⁹⁴. However, in the second data block, possibly due to the occurrence of habituation in the first block for both groups, no significant difference is observed in the trend of the slope of complexity between the two groups (see Fig. 5 for the trend of complexity slope over all channels and Fig. 6 for the trend of complexity slope over different brain regions). In mathematical terms, this change in the dynamics of complexity in brain activities can provide a new explanation for habit formation disorders^{55–57} and habituation disorders^{53,54} through interpreting them as aberrations in encoding and integration of information during the learning process to form a conditional distribution associated with the desired behavior.

Building on these observations, our findings corroborate habituation deficits in PD patients through differentiated complexity dynamics compared to the control group, offering a novel computational perspective on impaired learning dynamics related to the processing of uncertainty. Specifically, EEG complexity in the control group exhibited habituation properties known to occur during learning: a decreasing trend during sequence learning (Figs. 4–6), reflecting response attenuation⁴⁰, and spontaneous recovery, represented by a significantly higher complexity at the start of block 2 compared to the end of block 1 (Fig. 7). In PD patients, complexity convergence was delayed in block 1 compared to controls, particularly in the frontal region ($p < 0.001$, Fig. 6a), suggesting impaired uncertainty reduction. This aligns with reported PD habituation deficits^{53,54}, linked to slower cognitive processing^{79,80}, and striatal dysfunction^{57,94} which impairs information integration critical for sequence learning and habitual control^{2,64,91–93}. Based on these findings, we posit that delayed complexity convergence in PD reflects aberrations in encoding and integrating information to form conditional distributions during learning, contributing to habituation deficit reported in PD. However, the modest correlation between complexity and uncertainty (median ≈ 0.2) and challenges in EEG-based complexity modeling, as detailed in the section on Limitations of study, suggest that complexity may be one of the several factors which these deficits are reflected in, and that our findings may not imply causality. Despite these limitations, significant differences between the PD and control groups supported by robust statistical tests and aligned with prior PD literature, suggest complexity as a marker of habituation impairments in PD, warranting further investigation into underlying mechanisms.

Limitations of study

This study has several limitations that contextualize the scope of its findings. A primary limitation is the modest correlation values between EEG complexity and uncertainty (median ≈ 0.2), which suggest that uncertainty may explain only a small portion of the variance in complexity. While these correlations are statistically significant compared to chance level and are also significantly superior to those with other surprise metrics (Figs. 1–3, Table 1), their magnitudes do not indicate a strong relation between uncertainty and brain complexity. This modest effect size, although typical in EEG studies of cognitive processes, suggests that uncertainty processing in sequence learning may only be one of the several factors contributing to brain complexity. Additionally, the EEG data's limited spatial resolution hinders precise localization of the neural sources underlying the complexity–uncertainty relationship, particularly in distinguishing subcortical contributions such as the effects of striatal dysfunction in PD. Furthermore, our findings are associative and do not establish causality for habituation deficits in PD patients. For example, the observed delayed complexity convergence in the PD group (Figs. 4–6) aligns with known learning deficits of PD^{53,54} but does not allude to any causal mechanisms. Similarly, the data do not support a mechanistic view of how complexity represents uncertainty, as the significant correlations and convergence trends presented in our results (Figs. 1–6) are observational, not explanatory. These limitations highlight the need for future studies with higher-resolution neuroimaging (e.g., fMRI) to localize effects, longitudinal designs to probe causality, and multimodal approaches to elucidate the mechanisms of uncertainty processing and habituation impairments in PD. Despite these constraints, the robust statistical evidence (Figs. 1–7) and the alignment of our observations with PD and striatum literature^{53,54,79,80,91–94} provide a foundation for understanding the role of the brain's complexity and its fluctuations in sequence learning and PD-related deficits.

Conclusion

In examining the relationship between complexity and the learning process, our findings posit the idea that complexity of the brain activities, interpreted as the degree of information integration, reflects the learner's uncertainty. This uncertainty is formulated as the entropy of the conditional distribution of potential observations given the internal model shaped during the learning process. In the case of sequence learning, the conditional distribution can be considered as the distribution formed for the stimulus space given the observed stimuli.

In other words, given that the entropy of a distribution can mathematically be defined as the average code length for a random variable with that distribution, we can view the complexity of brain activities as an indicator of the amount of information from a new stimulus that is encoded and integrated with information from previous stimuli. This complexity therefore serves as a manifest of the process of updating the current belief of the learner over the stimuli space. Furthermore, the significant superiority of the correlation of brain complexity with the entropy of the posterior distribution formed through the learning process compared to its correlation with the Shannon and Bayesian models of surprise could indicate that complexity better represents uncertainty about the entire process of sequence generation and all stimuli, compared to just the novel information content in the recently observed stimulus as measured by the Shannon or Bayesian models of input surprise.

As a general description of the learning process, at the initial stage, a learner has a high level of uncertainty and needs to encode more information and integrate it with the current knowledge to expand their understanding of the situation. But as more information is gathered, the level of the learner's uncertainty may be reduced and there may be no need to encode or integrate as much information with the existing knowledge. Therefore, due to this general perspective and the notion of an association between the complexity of brain activities and uncertainty, which is mathematically termed as the amount of information coded by the learner, we expect that during the learning process, the complexity of the brain activities will show a decreasing trend, eventually converging to an asymptotic value. Our findings are consistent with this expectation and confirm a decreasing trend in brain complexity for the CTL group.

Additionally, this convergent behavior in brain complexity can be associated with the habituation process or habit formation, in which forming a habit can be likened to establishing a conditional distribution. Intuitively, forming a conditional distribution during a learning process such as habit formation can be viewed as a process

of achieving a certain level of certainty about the phenomenon, which is reflected in a convergent trend in the complexity of the brain activity.

Furthermore, due to the association between complexity and the habituation process, we expect to see some disparity in the trends of complexity of the brain activity between PD patients and CTL subjects due to the habituation and habit formation disorders caused by PD. Consistent with this expectation, a delayed convergence was observed in the complexity trend for PD patients compared to CTL subjects, corroborating the delay in cognitive processing and habituation and habit formation impairments attributed to PD.

As future work, the dynamic pattern of brain activity complexity can be monitored in setups where reward processing is involved or goal-directed behavior is present. Such studies can offer further insight into the behavior of complexity trends in the presence of a goal or a reward in the learning process. Furthermore, our analysis can be extended to invasive data recorded from animals at specific sites within the striatum and other regions of the basal ganglia during the processes of learning and habit formation. Such extension may allow for examining complexity signatures within these regions and following its trends locally as integrating new information progresses during the learning and habituation processes. Additionally, studying the dynamics of the brain activity complexity during other cognitive tasks that involve varied levels of attention can be the subject of further research. Furthermore, probing of the relationship between brain complexity and instantaneous attention level and any causal links between them can shed light on some of the intricate mechanisms that are involved in consciousness.

Methods

Dataset

The dataset used in this study is an open-access dataset, originally compiled and released in⁵⁴. Detailed methods of data acquisition and an in-depth explanation of the experimental tasks can be found in⁵⁴, but a brief description is provided here.

Twenty-five individuals with PD ($N = 25$) and an equal number of matched controls, of similar sex and age, (CTL group), participated in the experiment. The task was a 3-stimulus auditory oddball task. Participants were presented with a sequence composed of three sounds in a random order which was fixed for all participants: standard tones at a 440 Hz sinusoidal frequency (70% of trials), target tones at a 660 Hz sinusoidal frequency (15% of trials), and novelty distractors sourced from a naturalistic sound dataset (15% of trials). All sounds lasted for 200 ms. The inter-trial interval (ITI) was randomly chosen from a uniform distribution between 500 to 1000 ms. Single tone sounds were played at 80 dB, while novelty sounds were played at an average of 65 dB. Participants were instructed to mentally count the number of target stimuli. The task consisted of two blocks of 100 trials with similar sequential patterns of stimuli. A short break period occurred between the two blocks during which the participants reported their count of target stimuli.

Throughout the task, EEG signals were recorded using the standard 10–20 system. The signals were sampled at 500 samples per second across 64 channels, with CPz chosen as the online reference and AFz as the ground. Vertical electro-oculogram (VEOG) was also recorded from bipolar auxiliary inputs.

Analysis pipeline

To explore the relationship between complexity derived from brain activities and information-theoretical measures from Bayesian learning methods, the procedure depicted in Fig. 8 was implemented. EEG data were recorded during two blocks of an auditory oddball task consisting of three types of stimuli as described in Fig. 8. Following the preprocessing of the EEG data, measures of complexity were calculated over the time window of 200 ms to 400 ms post-stimulus onset for each trial. Complexity was measured using two established methods: Lempel–Ziv complexity (LZC) and sample entropy (SampEn). From the Bayesian learner's perspective, information-theoretical measures of surprise, namely Shannon surprise, KL divergence on the parameter and stimulus spaces, and uncertainty were associated to each stimulus presented in a trial. The correlations of each surprise metric with the measures of brain activity complexity were calculated over all trials. The maximum value for each correlation for a subject was chosen by tuning the parameter w which models the memory length of the subject. Furthermore, trend analysis was applied to the brain complexity measures throughout the trials to assess their behavior during the learning process. The correlation and trend analysis procedures are described in the following sections.

Preprocessing

All EEG data processing was performed using the EEGLab toolbox through an automated procedure. Initially, all non-EEG channels, including the EOG channel, were removed, leaving 63 channels. For each channel, the mean of the EEG data was subtracted to eliminate the direct current offset. A bandpass filter with a low cut-off frequency of 1 Hz and a high cut-off frequency of 150 Hz was then applied. This was followed by a 60 Hz notch filter with a 0.2 Hz width. Channels were removed if they exhibited any of the following attributes:

- 1) A flat period lasting at least 5 s,
- 2) Line noise to signal ratio (measured as the ratio of standard deviation to the grand mean of the channel) of 12 times or greater,
- 3) A correlation with nearby channels of less than 0.8.

Channels removed in the above step were interpolated using the spherical method⁹⁵. After interpolation, all EEG channel signals, including the interpolated ones, were re-referenced to the average of all signals. The Artifact Subspace Reconstruction (ASR) algorithm⁹⁶ was then applied to remove artifacts. The maximum acceptable standard deviation for the ASR algorithm in EEGLab, set for removing a bad window of 0.5 s, was set at 20 based

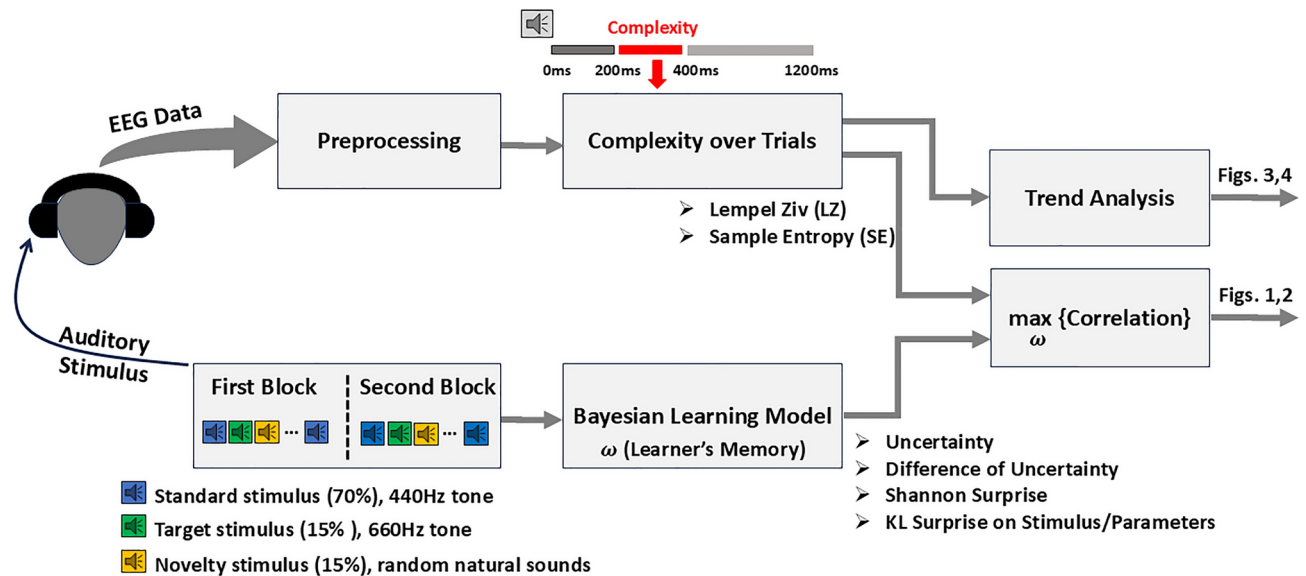


Fig. 8. The overall analysis pipeline. The oddball task consists of two blocks, each containing three different types of stimuli: standard, target, and novelty, with respective occurrence frequencies of 70%, 15% and 15%. The analysis is performed in two paths: EEG data analysis and information-theoretical surprise analysis. In the EEG data analysis path, the complexity of the EEG time series is calculated for each channel during the 200 ms to 400 ms time window after the stimulus onset for each trial. Furthermore, the trend of complexity throughout all trials is analyzed, see Figs. 4, 5, and 6. In the surprise analysis path, a Bayesian learning method with a parameter ω , capturing learner's memory, is applied to the sequence of stimuli presented to each subject. Various information-theoretical metrics are calculated and associated to each trial, capturing uncertainty, surprise, or the extent of the internal model update through the Bayesian learning process. Then, each of these Bayesian learning measures is optimized for each subject by selecting the value of the parameter ω which yields the highest correlation with the brain activity complexity. The maximum correlation values in each case are used to represent each subject and comparisons are made between the correlation values thereby obtained, see Figs. 1, 2, and 3.

on a visual check of preprocessed data from several sessions. Finally, channels were re-referenced to the average EEG channel. All preprocessing steps were visually validated on data from several participants to ensure their efficacy. These steps were adapted from Makoto's preprocessing pipeline⁹⁷.

Lempel–Ziv complexity (LZC) and sample entropy (SampEn)

Lempel–Ziv complexity is an algorithmic measure of the information content in a time series. The algorithm estimates the complexity of a sequence of symbols by counting the number of new patterns generated throughout the sequence. We implemented this algorithm as described in¹¹ to measure the complexity of EEG data. Other studies have employed this algorithm for complexity analysis in EEG⁹⁸ and magnetoencephalography (MEG)⁹⁹ data.

However, while Lempel–Ziv complexity has been extensively used in many studies and outperforms many other measures defined for complexity¹⁰⁰, it does not employ all information in the input data as it first quantizes the input time series into two binary levels¹⁰¹. Therefore, we also implemented the sample entropy method⁶ to assess the validity of our results based on a more fine-grained use of the recorded brain data.

To analyze EEG data complexity in the oddball task, EEG epochs were extracted between 200 and 400 ms after each stimulus onset for each channel, and both complexity methods were applied to each epoch. The average of each complexity measure over all channels was then calculated for each epoch and used in further analyses.

Transition probability model and surprise measures

We used an ideal observer model for the brain in its inference of the oddball input^{3,30}. This model assumes that the inferred knowledge of the brain about a sequence can be represented by a set of transition probabilities between the different stimuli types. In other words, during sequence learning, the brain estimates the probabilities of transition θ , an $n \times n$ matrix ($n=3$ is the number of stimulus types in the oddball task), in which θ_{ij} indicates the conditional probability of the occurrence of stimulus type j immediately after the observation of stimulus type i .

Assuming that $S_{1:t}$ indicates the sequence of stimuli having occurred from the first trial to the current trial, we can calculate the probability of $S_{1:t}$ as follows:

$$P(S_{1:t}) = P(S_1) \prod_{i=2}^t P(S_i | S_{i-1}, \theta)$$

which can be rewritten as:

$$P(S_{1:t}) = \prod_{i=1}^n \left(\prod_{j=1, j \neq i}^n \theta_{j|i}^{N_{j|i}} \right) \left(1 - \sum_{j=1, j \neq i}^n \theta_{j|i} \right)^{N_{j|i}}$$

In this equation, $N_{j|i}$ indicates the number of transitions from stimulus i to stimulus j in $S_{1:t}$. The posterior distribution of the transition probability parameter space can be calculated based on the Bayes rule as:

$$P(\theta|S_{1:t}) = \text{Dir}(\theta_1|N_{1|1}^t, N_{2|1}^t, N_{3|1}^t) \text{Dir}(\theta_2|N_{1|2}^t, N_{2|2}^t, N_{3|2}^t) \text{Dir}(\theta_3|N_{1|3}^t, N_{2|3}^t, N_{3|3}^t),$$

where θ_1, θ_2 , and θ_3 are rows of θ , and Dir indicates Dirichlet distribution. Thus, we can write the probability of S_t knowing $S_{1:t-1}$ as follows:

$$P(S_t|S_{1:t-1}) = \int P(S_t|S_{1:t-1}, \theta) P(\theta|S_{1:t-1}) d\theta$$

Based on the definition of θ we can write:

$$P(S_t|S_{1:t-1}, \theta) = P(S_t|S_{t-1}, \theta) = \theta_{ij}, S_t = j, S_{t-1} = i$$

On the other hand, the limitation of memory for each subject can be modeled as follows¹:

$$N_{j|i}^t = \sum_{k=2}^{t-1} \exp\left(-\frac{(t-k-1)}{w}\right) I\{S_k = i, S_{k-1} = j\}$$

where $I\{S_k = i, S_{k-1} = j\}$ is an indicator function which equals to one when the condition $S_k = i, S_{k-1} = j$ is satisfied and to zero otherwise, and $N_{j|i}^t$ is the number of transitions from state i to state j , and can be written in a recursive fashion as:

$$N_{j|i}^t = \exp(-1/w) N_{j|i}^{t-1} + I\{S_k = i, S_{(k-1)} = j\},$$

in which w is the parameter representing memory. Different measures of surprise can be defined based on this model. First, Shannon surprise is defined as follows:

$$\text{Shannon surprise}(t) = -\log(P(S_t|S_{1:t-1}))$$

Shannon surprise can be considered as a representation of puzzlement surprise³⁴. In the category of enlightenment surprises³⁴, the Kullback–Leibler divergence, introduced by^{83,84} and mathematically formulated to represent the mean difference between the prior and posterior distributions of the parameter space before and after observing stimulus S_t , is written as:

$$KL \text{ on transition probability}(t) = KL(P(\theta|S_{1:t})||P(\theta|S_{1:t-1}))$$

Extending the idea of enlightenment surprise as the amount of update in the internal model, we can also consider the Kullback–Leibler divergence between the prior and posterior distributions of the stimulus space before and after observing stimulus S_t :

$$KL \text{ on stimulus probability}(t) = KL(P(S_{t+1}|S_{1:t})||P(S_t|S_{1:t-1}))$$

In addition, uncertainty, representing the average surprise weighted by the learned conditional distribution after observing stimulus S_t , can be defined as:

$$\text{Uncertainty}(t) = H(P(S_{t+1}|S_{1:t})) = \sum_{i=1}^n -P(S_{t+1} = i|S_{1:t}) \log(P(S_{t+1} = i|S_{1:t})),$$

where H refers to the entropy of the conditional distribution of the next stimulus given the previous stimuli (including the current observed stimulus). In fact, uncertainty represents the average Shannon surprise which the subject expects to experience. Finally, difference-of-uncertainty, as a representation of enlightenment surprise, is calculated as: $\text{Diff-Uncertainty}(t) = \text{Uncertainty}(t) - \text{Uncertainty}(t-1)$

Correlation calculation

Applying each measure of surprise to the input sequence yields a sequence of values derived from the Bayesian learning model as described above. On the other hand, for each subject a sequence of complexity values corresponding to the stimulus sequence is derived from the subject's EEG data. Given the potential differences in memory capabilities among subjects, an adjustment over the memory length parameter w may be necessary for each individual. Thus, the correlation between the complexity sequence and the various surprise measure

sequences was optimized over w for each subject. All subsequent analyses of correlation between measures of surprise and brain complexity were based on the optimized correlation value for each individual.

To assess the significance of these correlations, we implemented two hypothesis tests. In the first approach, called the P1 test, the complexity sequence for each subject was subjected to 200 random permutations. For each permutation, the correlation value was calculated and optimized over the parameter w for each subject. In the second approach, called the P2 test, the stimulus sequence for each subject underwent 200 random permutations. Subsequently, the surprise sequence was recalculated and the correlations were determined and optimized over the parameter w for each subject. For each of the two tests, to assess the hypotheses that each measure of surprise or uncertainty had a higher correlation with brain complexity than its corresponding surrogate correlation, the original set of 25 correlations were compared with the correlations resulting from the permutations—totaling 200 (number of permutations) times 25 (number of subjects in each group)—using a one-tailed t-test with unequal variance.

To compare the extent of relationship of the Bayesian information metrics with the complexity of brain activities, we categorized these metrics into two groups of enlightenment and puzzlement surprises. The puzzlement surprise group refers to the two metrics of uncertainty and Shannon surprise. For this group, we compared their correlations with the brain complexity over trials using paired t-test analysis. The enlightenment surprise group consists of the KL-divergence measures corresponding to the parameter and stimulus spaces, as well as difference-of-uncertainty. For this group, we compared their correlations with the change in brain complexity from one trial to the next, which reflects the change in brain activities upon observing the current stimulus. This comparison was also conducted using paired t-test analysis.

Trend analysis of complexity in oddball task

To study the complexity trends, we applied a moving average window with a length of 30 trials to the complexity values across trials for each subject. This served as a low-pass filter to smooth out fluctuations in complexity across trials. Additionally, to study habituation and the decline in complexity values over time, a least-squares line was fitted to the set of consecutive complexity values in a moving window of 40 samples. The trend of the slope of the fitted lines was also calculated over the sequence of stimuli for each subject.

Statistical analyses

To compare the degree of association between the various information metrics from Bayesian learning and the complexity of brain activities, paired t-test analyses were used. Each test compared the correlation of a pair of information metrics with a measure of brain complexity, under the null hypothesis that both information metrics correlate similarly with brain complexity on average for each subject group.

Additionally, to determine if information metrics significantly correlate with brain complexity, two distinct permutations were applied: one to the stimuli and another to complexity values. For the first permutation, 200 permutations of the stimuli order in one block of the oddball task were executed, and all procedures, including optimization over the parameter w , were repeated to recalculate each information metric's correlation with brain complexity. This yielded 5000 correlation values (25 subjects $\times$ 200 permutations). Then, for each information metric, the original correlation values were compared to the recalculated (surrogate) correlations using a one-sided t-test with unequal variance, to ascertain whether the original correlations were significantly higher than the surrogate ones corresponding to the permutations. A similar procedure was employed for the second permutation method which was performed on the complexity values. Furthermore, the permutation analysis was separately executed on each EEG channel, resulting in topographic maps representing the resulting p-values, on which channels exhibiting significant correlations after Bonferroni correction could be identified.

To compare the dynamics of complexity values over trials between the PD and CTL groups, moving averages of the value and the slope of complexity were calculated for each subject. This resulted in two arrays indicating the trends of complexity and its slope across trials. Then, to compare these trends between the PD and CTL groups, the indices of each of these arrays were divided into three stages of early, middle, and late. Two-sided t-tests were performed on the average values within each stage for an array, yielding a p-value for each stage, leading to three p-values for each region.

Finally, to investigate the phenomenon of spontaneous recovery in complexity, which stipulates a higher level of complexity occurs at the beginning of the second data block compared to the end of the first block, a two-sided paired t-test for 25 paired values ($df = 24$) was employed to compare the average of the last ten complexity values in the first block with the average of the first ten complexities in the second block.

Data availability

The dataset analyzed in this study was produced by authors of an article published earlier and information about accessing the data can be found within that article: J. F. Cavanagh, P. Kumar, A. A. Mueller, S. P. Richardson, and A. Mueen, “Diminished EEG habituation to novel events effectively classifies Parkinson's patients,” *Clin Neurophysiol*, vol. 129, no. 2, pp. 409–418, Feb. 2018, <https://doi.org/10.1016/j.clinph.2017.11.023>.

Received: 30 May 2024; Accepted: 29 April 2025

Published online: 01 July 2025

References

1. J. Von Neumann, *The computer and the brain*. In the computer and the brain. Oxford England: Yale Univer. Press xiv, 82 1958
2. Tononi, G., Edelman, G. M. & Sporns, O. Complexity and coherency: Integrating information in the brain. *Trends Cogn. Sci.* 2(12), 474–484. [https://doi.org/10.1016/S1364-6613\(98\)01259-5](https://doi.org/10.1016/S1364-6613(98)01259-5) (1998).

3. Meyniel, F., Maheu, M. & Dehaene, S. Human inferences about sequences: A minimal transition probability model. *PLoS Comput. Biol.* **12**(12), e1005260. <https://doi.org/10.1371/journal.pcbi.1005260> (2016).
4. Sporns, O. The complex brain: Connectivity, dynamics, information. *Trends Cogn. Sci.* **26**(12), 1066–1067. <https://doi.org/10.1016/j.tics.2022.08.002> (2022).
5. Bassett, D. S. & Gazzaniga, M. S. Understanding complexity in the human brain. *Trends Cogn. Sci.* **15**(5), 200–209. <https://doi.org/10.1016/j.tics.2011.03.006> (2011).
6. Shorten, D. P., Priesemann, V., Wibral, M. & Lizier, J. T. Early lock-in of structured and specialised information flows during neural development. *Elife* **11**, e74651. <https://doi.org/10.7554/eLife.74651> (2022).
7. Chén, O. Y. et al. Resting-state brain information flow predicts cognitive flexibility in humans. *Sci. Rep.* <https://doi.org/10.1038/s41598-019-40345-8> (2019).
8. Tian, X. et al. An integrated resource for functional and structural connectivity of the marmoset brain. *Nat. Commun.* **13**(1), 7416. <https://doi.org/10.1038/s41467-022-35197-2> (2022).
9. Jiang, J., Wagner, A. D. & Egner, T. Integrated externally and internally generated task predictions jointly guide cognitive control in prefrontal cortex. *Elife* **7**, e39497. <https://doi.org/10.7554/eLife.39497> (2018).
10. Chung, A. et al. Cognitive control persistently enhances hippocampal information processing. *Nature* <https://doi.org/10.1038/s41586-021-04070-5> (2021).
11. Lempel, A. & Ziv, J. On the complexity of finite sequences. *IEEE Trans. Inf. Theory* **22**(1), 75–81. <https://doi.org/10.1109/TIT.1976.1055501> (1976).
12. Richman, J. S. & Moorman, J. R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* **278**(6), H2039–2049. <https://doi.org/10.1152/ajpheart.2000.278.6.H2039> (2000).
13. Delgado-Bonal, A. & Marshak, A. Approximate entropy and sample entropy: A comprehensive tutorial. *Entropy (Basel)* **21**(6), 541. <https://doi.org/10.3390/e21060541> (2019).
14. Modis, T. Links between entropy, complexity, and the technological singularity. *Technol. Forecast. & Soc. Change* **176**, 121457. <https://doi.org/10.1016/j.techfore.2021.121457> (2022).
15. Teixeira, A., Matos, A., Souto, A. & Antunes, L. Entropy measures vs. kolmogorov complexity. *Entropy* <https://doi.org/10.3390/e13030595> (2011).
16. A. Kaltchenko, Algorithms for estimating information distance with application to bioinformatics and linguistics In *Canadian Conference on Electrical and Computer Engineering 2004 (IEEE Cat. No.04CH37513)* 2255–2258 4 <https://doi.org/10.1109/CCEC.E.2004.1347695> 2004
17. Li, M. & Vitányi, P. *An Introduction to Kolmogorov Complexity and Its Applications*. <https://doi.org/10.1007/978-1-4757-2606-0> (1997).
18. Hsu, C. F. et al. Entropy of entropy: Measurement of dynamical complexity for biological systems. *Entropy* <https://doi.org/10.3390/e19100550> (2017).
19. J. Amigó, *Permutation complexity in dynamical systems: Ordinal patterns, permutation entropy and all that*. Springer Science spsamps Business Media (2010).
20. *Entropy of Complex Processes and Systems* by Eugene Barsky - Ebook | Everand. Accessed: Dec. 04, 2023. [Online]. Available: <https://www.everand.com/book/469998545/Entropy-of-Complex-Processes-and-Systems>
21. Nastase, S. A., Iacovella, V., Davis, B. & Hasson, U. Connectivity in the human brain dissociates entropy and complexity of auditory inputs. *Neuroimage* **108**, 292–300. <https://doi.org/10.1016/j.neuroimage.2014.12.048> (2015).
22. Vecchio, F. et al. Entropy as measure of brain networks' complexity in eyes open and closed conditions. *Symmetry* <https://doi.org/10.3390/sym13112178> (2021).
23. Keshmiri, S. Entropy and the brain: An overview. *Entropy (Basel)* **22**(9), 917. <https://doi.org/10.3390/e22090917> (2020).
24. Mateos, D. M., Guevara Erra, R., Wennberg, R. & Perez Velazquez, J. L. Measures of entropy and complexity in altered states of consciousness. *Cogn. Neurodynamics* **12**(1), 73–84. <https://doi.org/10.1007/s11571-017-9459-8> (2018).
25. Sarasso, S. et al. Consciousness and complexity: a consilience of evidence. *Neurosci. Conscious.* <https://doi.org/10.1093/nc/niab023> (2021).
26. Crumiller, M., Knight, B., Yu, Y. & Kaplan, E. Estimating the amount of information conveyed by a population of neurons. *Front. Neurosci.* **5**, 90. <https://doi.org/10.3389/fnins.2011.00090> (2011).
27. H. Feldman and K. Friston, “Attention, Uncertainty, and Free-Energy,” *Frontiers in Human Neuroscience* 4 Accessed: Nov. 14, 2022. [Online]. Available: <https://www.frontiersin.org/articles/https://doi.org/10.3389/fnhum.2010.002152010>
28. Saffran, J. R., Aslin, R. N. & Newport, E. L. Statistical learning by 8-month-old infants. *Science* **274**(5294), 1926–1928. <https://doi.org/10.1126/science.274.5294.1926> (1996).
29. Takacs, A. & Beste, C. A neurophysiological perspective on the integration between incidental learning and cognitive control. *Commun. Biol.* <https://doi.org/10.1038/s42003-023-04692-7> (2023).
30. Maheu, M., Dehaene, S. & Meyniel, F. Brain signatures of a multiscale process of sequence learning in humans. *Elife* **8**, e41541. <https://doi.org/10.7554/eLife.41541> (2019).
31. Runyan, C. A., Piasini, E., Panzeri, S. & Harvey, C. D. Distinct timescales of population coding across cortex. *Nature* <https://doi.org/10.1038/nature23020> (2017).
32. Yang, Y., DeWeese, M. R., Otazu, G. H. & Zador, A. M. Millisecond-scale differences in neural activity in auditory cortex can drive decisions. *Nat. Neurosci.* <https://doi.org/10.1038/nn.2211> (2008).
33. Hasson, U., Chen, J. & Honey, C. J. Hierarchical process memory: Memory as an integral component of information processing. *Trends Cogn. Sci.* **19**(6), 304. <https://doi.org/10.1016/j.tics.2015.04.006> (2015).
34. Faraji, M., Preuschoff, K. & Gerstner, W. Balancing new against old information: The role of puzzlement surprise in learning. *Neural. Comput.* **30**(1), 34–83. https://doi.org/10.1162/neco_a_01025 (2018).
35. Jeong, W., Kim, S., Park, J. & Lee, J. Multivariate EEG activity reflects the Bayesian integration and the integrated Galilean relative velocity of sensory motion during sensorimotor behavior. *Commun. Biol.* <https://doi.org/10.1038/s42003-023-04481-2> (2023).
36. P. Juslin, Controlled information integration and bayesian inference *Frontiers in Psychology* 6 Accessed: May 29, 2023. [Online]. Available: <https://www.frontiersin.org/articles/https://doi.org/10.3389/fpsyg.2015.00070> (2015)
37. Balduzzi, D. & Tononi, G. Integrated information in discrete dynamical systems: motivation and theoretical framework. *PLoS Comput. Biol.* **4**(6), e1000091. <https://doi.org/10.1371/journal.pcbi.1000091> (2008).
38. Friston, K. The free-energy principle: a unified brain theory?. *Nat. Rev. Neurosci.* **11**(2), 127–138. <https://doi.org/10.1038/nrn2787> (2010).
39. Friston, K. Hierarchical models in the brain. *PLoS Comput. Biol.* **4**(11), e1000211. <https://doi.org/10.1371/journal.pcbi.1000211> (2008).
40. Rankin, C. H. et al. Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. *Neurobiol. Learn Mem.* **92**(2), 135–138. <https://doi.org/10.1016/j.nlm.2008.09.012> (2009).
41. Thompson, R. F. & Spencer, W. A. Habituation: a model phenomenon for the study of neuronal substrates of behavior. *Psychol. Rev.* **73**(1), 16–43. <https://doi.org/10.1037/h0022681> (1966).
42. Tononi, G. An information integration theory of consciousness. *BMC Neurosci.* **5**(1), 42. <https://doi.org/10.1186/1471-2202-5-42> (2004).
43. Tononi, G. & Edelman, G. M. Consciousness and complexity. *Science* **282**(5395), 1846–1851. <https://doi.org/10.1126/science.282.5395.1846> (1998).

44. Grossberg, S. The link between brain learning, attention, and consciousness. *Conscious Cogn.* **8**(1), 1–44. <https://doi.org/10.1006/ccog.1998.0372> (1999).
45. Lisman, J. & Sternberg, E. J. Habit and nonhabit systems for unconscious and conscious behavior: implications for multitasking. *J. Cogn. Neurosci.* **25**(2), 273–283. https://doi.org/10.1162/jocn_a.00319 (2013).
46. Dauer, W. & Przedborski, S. Parkinson's disease: Mechanisms and models. *Neuron* **39**(6), 889–909. [https://doi.org/10.1016/S0896-6273\(03\)00568-3](https://doi.org/10.1016/S0896-6273(03)00568-3) (2003).
47. Albin, R. L., Young, A. B. & Penney, J. B. The functional anatomy of basal ganglia disorders. *Trends Neurosci.* **12**(10), 366–375. [https://doi.org/10.1016/0166-2236\(89\)90074-x](https://doi.org/10.1016/0166-2236(89)90074-x) (1989).
48. Cools, R. Dopaminergic modulation of cognitive function-implications for L-DOPA treatment in Parkinson's disease. *Neurosci. Biobehav. Rev.* **30**(1), 1–23. <https://doi.org/10.1016/j.neubiorev.2005.03.024> (2006).
49. Brown, R. G. & Marsden, C. D. Internal versus external cues and the control of attention in Parkinson's disease. *Brain* **111**(Pt 2), 323–345. <https://doi.org/10.1093/brain/111.2.323> (1988).
50. Dujardin, K., Degre, J. F., Rogelet, P., Defebvre, L. & Destee, A. Impairment of the supervisory attentional system in early untreated patients with Parkinson's disease. *J. Neurol.* **246**(9), 783–788. <https://doi.org/10.1007/s004150050455> (1999).
51. Cools, R., Rogers, R., Barker, R. A. & Robbins, T. W. Top-down attentional control in Parkinson's disease: Salient considerations. *J. Cogn. Neurosci.* **22**(5), 848–859. <https://doi.org/10.1162/jocn.2009.21227> (2010).
52. Bocquillon, P. et al. Role of basal ganglia circuits in resisting interference by distracters: A swLORETA study. *PLoS ONE* **7**(3), e34239. <https://doi.org/10.1371/journal.pone.0034239> (2012).
53. Jiang, C., Kaseda, Y., Kumagai, R., Nakano, Y. & Nakamura, S. Habituation of event-related potentials in patients with Parkinson's disease. *Physiol. Behav.* **68**(5), 741–747. [https://doi.org/10.1016/S0031-9384\(99\)00244-9](https://doi.org/10.1016/S0031-9384(99)00244-9) (2000).
54. Cavanagh, J. F., Kumar, P., Mueller, A. A., Richardson, S. P. & Mueen, A. Diminished EEG habituation to novel events effectively classifies parkinson's Patients. *Clin. Neurophysiol.* **129**(2), 409–418. <https://doi.org/10.1016/j.clinph.2017.11.023> (2018).
55. Mi, T.-M., Zhang, W., McKeown, M. J. & Chan, P. Impaired formation and expression of goal-directed and habitual control in parkinson's disease. *Front. Aging. Neurosci.* **13**, 734807. <https://doi.org/10.3389/fnagi.2021.734807> (2021).
56. Marzinzik, F. et al. Modulation of habit formation by levodopa in Parkinson's Disease. *PLoS ONE* **6**(11), e27695. <https://doi.org/10.1371/journal.pone.0027695> (2011).
57. Redgrave, P. et al. Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease. *Nat. Rev. Neurosci.* **11**, 11. <https://doi.org/10.1038/nrn2915> (2010).
58. Wang, L. et al. Effect of interstimulus interval on visual P300 in Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry.* **67**(4), 497–503. <https://doi.org/10.1136/jnnp.67.4.497> (1999).
59. Goto, Y. & Grace, A. A. Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. *Nat. Neurosci.* <https://doi.org/10.1038/nn1471> (2005).
60. O'Donnell, P. & Grace, A. A. Synaptic interactions among excitatory afferents to nucleus accumbens neurons: hippocampal gating of prefrontal cortical input. *J. Neurosci.* **15**(5 Pt 1), 3622–3639. <https://doi.org/10.1523/JNEUROSCI.15-05-03622.1995> (1995).
61. Atallah, H. E., Frank, M. J. & O'Reilly, R. C. Hippocampus, cortex, and basal ganglia: Insights from computational models of complementary learning systems. *Neurobiol. Learn. Mem.* **82**(3), 253–267. <https://doi.org/10.1016/j.nlm.2004.06.004> (2004).
62. van Schouwenburg, M. R., den Ouden, H. E. M. & Cools, R. The human basal ganglia modulate frontal-posterior connectivity during attention shifting. *J. Neurosci.* **30**(29), 9910–9918. <https://doi.org/10.1523/JNEUROSCI.1111-10.2010> (2010).
63. Lawrence, A. D., Sahakian, B. J. & Robbins, T. W. Cognitive functions and corticostriatal circuits: Insights from Huntington's disease. *Trends Cogn. Sci.* **2**(10), 379–388. [https://doi.org/10.1016/S1364-6613\(98\)01231-5](https://doi.org/10.1016/S1364-6613(98)01231-5) (1998).
64. Yin, H. H. & Knowlton, B. J. The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* <https://doi.org/10.1038/nrn1919> (2006).
65. Bornstein, A. M. & Daw, N. D. Multiplicity of control in the basal ganglia: Computational roles of striatal subregions. *Curr. Opin. Neurobiol.* **21**(3), 374–380. <https://doi.org/10.1016/j.conb.2011.02.009> (2011).
66. Kang, J. et al. Primate ventral striatum maintains neural representations of the value of previously rewarded objects for habitual seeking. *Nat. Commun.* <https://doi.org/10.1038/s41467-021-22335-5> (2021).
67. Markowitz, J. E. & Datta, S. R. The striatum specifies the statistics of behavior. *Neuropsychopharmacology* <https://doi.org/10.1038/s41386-019-0493-6> (2019).
68. Santos, F. J., Oliveira, R. F., Jin, X. & Costa, R. M. Corticostriatal dynamics encode the refinement of specific behavioral variability during skill learning. *Elife* **4**, e09423. <https://doi.org/10.7554/eLife.09423> (2015).
69. Yartsev, M. M., Hanks, T. D., Yoon, A. M. & Brody, C. D. Causal contribution and dynamical encoding in the striatum during evidence accumulation. *Elife* **7**, e34929. <https://doi.org/10.7554/eLife.34929> (2018).
70. Kravitz, A. V. & Matikainen-Ankney, B. A. Motor control: Memory and motor control in the dorsal striatum. *Curr. Biol.* **30**(22), R1366–R1368. <https://doi.org/10.1016/j.cub.2020.09.018> (2020).
71. Bäckman, L. et al. Effects of working-memory training on striatal dopamine release. *Science* **333**(6043), 718. <https://doi.org/10.1126/science.1204978> (2011).
72. Aarts, E. et al. Striatal dopamine mediates the interface between motivational and cognitive control in humans: Evidence from genetic imaging. *Neuropsychopharmacol* <https://doi.org/10.1038/npp.2010.68> (2010).
73. Beste, C., Moll, C. K. E., Pötter-Nerger, M. & Münchau, A. Striatal microstructure and its relevance for cognitive control. *Trends Cogn. Sci.* **22**(9), 747–751. <https://doi.org/10.1016/j.tics.2018.06.007> (2018).
74. Nee, D. E. Integrative frontal-parietal dynamics supporting cognitive control. *Elife* **10**, e57244. <https://doi.org/10.7554/eLife.57244> (2021).
75. Sharma, K., Mangarosa, K., van Berckel, N., Giannakos, M. & Kostakos, V. Information flow and cognition affect each other: Evidence from digital learning. *Int. J. Hum. Comput. Stud.* **146**, 102549. <https://doi.org/10.1016/j.ijhcs.2020.102549> (2021).
76. Wan, W., Cui, X., Gao, Z. & Gu, Z. Frontal EEG-based multi-level attention states recognition using dynamical complexity and extreme gradient boosting. *Front Hum. Neurosci.* **15**, 673955. <https://doi.org/10.3389/fnhum.2021.673955> (2021).
77. D. S. Bassett, N. F. Wymbs, M. A. Porter, P. J. Mucha, J. M. Carlson, and S. T. Grafton, "Dynamic reconfiguration of human brain networks during learning," *Proc. Natl. Acad. Sci.* **108** 18 7641–7646 <https://doi.org/10.1073/pnas.1018985108>. 2011
78. Grant, S. G. N. Synapse molecular complexity and the plasticity behaviour problem. *Brain Neurosci. Adv.* **2**, 2398212818810685. <https://doi.org/10.1177/2398212818810685> (2018).
79. Jokinen, P. et al. Cognitive slowing in Parkinson's disease is related to frontostriatal dopaminergic dysfunction. *J. Neurol. Sci.* **329**(1), 23–28. <https://doi.org/10.1016/j.jns.2013.03.006> (2013).
80. Stelmach, G. E., Teasdale, N. & Phillips, J. Response initiation delays in Parkinson's disease patients. *Hum. Mov. Sci.* **11**(1), 37–45. [https://doi.org/10.1016/0167-9457\(92\)90048-G](https://doi.org/10.1016/0167-9457(92)90048-G) (1992).
81. Friston, K. A theory of cortical responses. *Philos. Trans. Royal Soc. B: Biol. Sci.* **360**(1456), 815–836. <https://doi.org/10.1098/rstb.2005.1622> (2005).
82. Shannon, C. E. A mathematical theory of communication. *Bell Syst. Techn. J.* **27**(3), 379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x> (1948).
83. Itti, L. & Baldi, P. Bayesian surprise attracts human attention. *Vision Res.* **49**(10), 1295–1306. <https://doi.org/10.1016/j.visres.2009.09.007> (2009).
84. J. Storck, S. Hochreiter, and J. Schmidhuber, Reinforcement Driven Information Acquisition In Non-Deterministic Environments *ICANN'95* 2 (1995).

85. Ma, W. J. & Jazayeri, M. Neural coding of uncertainty and probability. *Annu. Rev. Neurosci.* **37**, 205–220. <https://doi.org/10.1146/annurev-neuro-071013-014017> (2014).
86. Knill, D. C. & Pouget, A. The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends Neurosci.* **27**(12), 712–719. <https://doi.org/10.1016/j.tins.2004.10.007> (2004).
87. N. Barascud, M. T. Pearce, T. D. Griffiths, K. J. Friston, and M. Chait, “Brain responses in humans reveal ideal observer-like sensitivity to complex acoustic patterns,” *Proc. of the National Academy of Sciences*, 113 5 E616–E625 <https://doi.org/10.1073/pnas.1508523113> (2016).
88. T. A. Bekinschtein, S. Dehaene, B. Rohaut, F. Tadel, L. Cohen, and L. Naccache, Neural signature of the conscious processing of auditory regularities *Proc. of the National Academy of Sciences*, 106 (5) 1672–1677, <https://doi.org/10.1073/pnas.0809667106> (2009).
89. Bayley, P. J., Frascino, J. C. & Squire, L. R. Robust habit learning in the absence of awareness and independent of the medial temporal lobe. *Nature* <https://doi.org/10.1038/nature03857> (2005).
90. Seger, C. A. & Spiering, B. J. A critical review of habit learning and the basal ganglia. *Front Syst. Neurosci.* **5**, 66. <https://doi.org/10.3389/fnsys.2011.00066> (2011).
91. Wilkinson, L. & Jahanshahi, M. The striatum and probabilistic implicit sequence learning. *Brain Res.* **1137**(1), 117–130. <https://doi.org/10.1016/j.brainres.2006.12.051> (2007).
92. K. R. Gamble, T. J. Cummings Jr., S. E. Lo, P. T. Ghosh, J. H. Howard Jr., and D. V. Howard, Implicit sequence learning in people with Parkinson's disease *Frontiers in Human Neuroscience* 8 Accessed: Aug. 14, 2023. [Online]. Available: <https://www.frontiersin.org/articles/https://doi.org/10.3389/fnhum.2014.00563> (2014)
93. Winkelmeier, L. et al. Striatal hub of dynamic and stabilized prediction coding in forebrain networks for olfactory reinforcement learning. *Nat. Commun.* **13**, 1. <https://doi.org/10.1038/s41467-022-30978-1> (2022).
94. Voon, V. et al. The neurochemical substrates of habitual and goal-directed control. *Transl. Psychiatry*. <https://doi.org/10.1038/s41398-020-0762-5ss> (2020).
95. Freedman, W. Spherical spline interpolation—basic theory and computational aspects. *J. Comput. Appl. Math.* **11**(3), 367–375. [https://doi.org/10.1016/0377-0427\(84\)90011-6](https://doi.org/10.1016/0377-0427(84)90011-6) (1984).
96. Mullen, T. et al. Real-time modeling and 3D visualization of source dynamics and connectivity using wearable EEG. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2013**, 2184–2187. <https://doi.org/10.1109/EMBC.2013.6609968> (2013).
97. “Makoto's preprocessing pipeline - SCCN.” [Online]. Available: https://sccn.ucsd.edu/wiki/Makoto's_preprocessing_pipeline (accessed Oct. 30, 2022).
98. Zhang, X. S., Roy, R. J. & Jensen, E. W. EEG complexity as a measure of depth of anesthesia for patients. *IEEE Trans. Biomed. Eng.* **48**(12), 1424–1433. <https://doi.org/10.1109/10.966601> (2001).
99. Gómez, C. et al. Complexity analysis of resting-state MEG activity in early-stage Parkinson's disease patients. *Ann. Biomed. Eng.* **39**(12), 2935–2944. <https://doi.org/10.1007/s10439-011-0416-0> (2011).
100. D. Aiordachioaie and T. D. Popescu, Aspects of Time Series Analysis with Entropies and Complexity Measures In *2020 International Symposium on Electronics and Telecommunications (ISETC)* 2020 1–4. <https://doi.org/10.1109/ISETC50328.2020.9301111>.
101. Ibáñez-Molina, A. J., Iglesias-Parro, S., Soriano, M. F. & Aznarte, J. I. Multiscale Lempel-Ziv complexity for EEG measures. *Clin. Neurophysiol.* **126**(3), 541–548. <https://doi.org/10.1016/j.clinph.2014.07.012> (2015).

Acknowledgements

The authors wish to thank J. F. Cavanagh and colleagues for making their dataset publicly available.

Author contributions

MHHB, KH, and MRS conducted the study and analyzed the data. HA supervised the work. All authors wrote, edited, and approved the manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to H.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025