UNIVERSITY
OF OSLO

Julian Fuhrer

Implicit Encoding of Seemingly Unstructured Auditory Stimuli

Thesis Submitted for the Degree of Philosophiae Doctor

Department of Informatics
Faculty of Mathematics and Natural Sciences

RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion

2023

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 2588

ISSN 1501-7710

Print production: Graphics Center, University of Oslo.

FAUST.

Das Pergament, ist das der heil'ge Bronnen,
Woraus ein Trunk den Durst auf ewig stillt?
Erquickung hast du nicht gewonnen,
Wenn sie dir nicht aus eigner Seele quillt.

Abstract

Humans are biased to perceive patterns in random sensory signals. While there are numerous examples of this phenomenon termed apophenia comprising the perception of faces in clouds, recognition of words within noise, constellations or the gambler's fallacy, the exact underlying neurophysiological mechanisms are unknown.

Presumably at the core of this phenomenon are internal representations that map the immediate environment, made possible by automatically drawing on statistical structures in the unfolding sensory input. Identifying how and where this processing operates is a core question in research fields such as statistical learning or predictive processing. Specifically, the exact role of different involved brain regions remains elusive. Identified areas are widely spread across the cortex, situating lower as well as higher up in the cortical hierarchy (e.g., perceptual systems or prefrontal cortices). This suggests a dynamic network composed of distributed brain regions carrying out individual actions.

Utilising the high temporal and spectral precision of intracranial electroencephalography, this thesis identified that the brain automatically encodes temporal statistical structures between events when exposed to random acoustic stimuli. This encoding involved a network outside the auditory system, including hippocampal, frontal and temporal cortices.

To achieve this, the effective representation of neurophysiological signals through information-theoretical principles was investigated, conceiving a novel approach titled Encoded Information. This method allows quantifying differences between recorded responses and consequently offers to discriminate conditions of a study design. Based on five cortical activity types originating from intracranial electroencephalography and simulated neurophysiological recordings suggested approach was compared with well-established methods and examined through a parameter study. The proposed procedure performs equally efficiently or better than, for example, the conventional t-test. In addition, it allows statements about the sensitivity toward the cortical encoding of unexpected events. Further, it constituted a viable measure in the main analysis in which it was employed to represent neurophysiological responses evoked by random acoustic stimuli which were then compared to estimated statistical structures within that stimuli.

Overall the present work suggests that the brain continuously attempts to predict and provide structure from events in the environment, even when they are not behaviourally relevant and have no evident relation between them. Linking the theoretical frameworks of statistical learning and predictive processing, this thesis illuminates the underlying implicit brain mechanisms that can be crucial for rapidly encoding patterns or unexpected events in the environment. Further, the method of Encoded Information introduces an apt candidate to represent and investigate neurophysiological data through information-theoretical principles.

Preface

This thesis is submitted in partial fulfilment of the requirements for the degree of *Philosophiae Doctor* at the University of Oslo. The research presented was conducted at the University of Oslo under the supervision of Kyrre Glette, Alejandro Blenkmann and Tor Endestad. This work was partly supported by the Research Council of Norway (RCN) through its Centres of Excellence scheme project number 262762, RCN project number 240389 and 314925.

The thesis is a collection of three papers presented in chronological order of writing. The common theme is the investigation of neurophysiological recordings through information-theoretical principles. The papers are preceded by an introductory chapter that relates them to each other and chapters providing background information and motivation for the work. After presenting the papers, a discussion aiming to set the papers into a bigger context follows.

Julian Fuhrer

Oslo, January 2023

List of Papers

Paper I

Fuhrer, J., Glette, K., Ivanovic, J., Larsson, P. G., Bekinschtein, T., Kochen, S., Knight, R. T., Tørresen, J., Solbakk, A.-K., Endestad, T., & Blenkmann, A. (submitted). Direct brain recordings reveal continuous encoding of structure in random stimuli. *bioRxiv*. https://doi.org/10.1101/2021.10.01.462295

Paper II

Fuhrer, J., Blenkmann, A., Endestad, T., Solbakk, A.-K., & Glette, K. (2022). Complexity-based encoded information quantification in neurophysiological recordings. *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2319–2323. https://doi.org/10.1109/EMBC48229.2022.9871501

Paper III

Fuhrer, J., Glette, K., Endestad, T., Solbakk, A.-K., & Blenkmann, A. (submitted). Quantifying evoked responses through encoded information. *bioRxiv*. https://doi.org/10.1101/2022.11.11.516096

Contents

Chapter 1

Introduction

The auditory system is an early warning system. It is designed in order for us to survive. Responsible for this is automatic processing, even active during sleep or coma (Atienza et al., 1997; Bekinschtein et al., 2009; Fischer et al., 1999). Mounting evidence suggests that this processing operates through sophisticated internal perceptual models. Being continuously refined as the sensory input unfolds, they enable the brain to quickly discover and evaluate structure within the sensory input stream and to maintain a detailed representation of the environment (Bastos et al., 2015, 2020; Clark, 2013; Dürschmid et al., 2016; Friston, 2008, 2010; Paavilainen, 2013; Rao & Ballard, 1999; Saffran et al., 1996; Thiessen, 2017). When a sensory event inconsistent with the established model is encountered, a "surprise" response is produced, facilitating a swift reaction to the respective real cause. This encoding of environmental patterns is made possible by exploiting statistical structure in space and time within the sensory signal (Conway, 2020; Daikoku, 2018; Dehaene et al., 2015; Maheu et al., 2019; Saffran & Kirkham, 2018; Southwell & Chait, 2018; Thiessen, 2017).

Our perceptual systems have been confronted with repeating sensory signals over thousand of years (Turk-Browne, 2012). Hence these systems were able to adapt to repeated aspects of the environment, named regularities. Studies have shown that starting from an early age we are sensitive to two kinds of statistical regularities, namely distributional and conditional regularities. Distributional regularities describe the frequency or variability of event types and are presumably exploited to group similar events into categories. Conditional regularities, on the other hand, relate different event types to each other and are likely a core statistical structure used for grouping events into larger elements (e.g., syllables to words). Importantly, as the sensory input stream unfolds, the brain makes use of both structures (Baumgarten et al., 2021; Dehaene et al., 2015; Domenech & Dreher, 2010; Henin et al., 2021; Higashi et al., 2017; Koelsch et al., 2016; Leonard et al., 2015; Maguire et al., 2019; Maheu et al., 2022; Meyniel & Dehaene, 2017; Meyniel et al., 2015, 2016; Mittag et al., 2016; Saffran, 2020; Thiessen, 2017; Turk-Browne, 2012; Turk-Browne et al., 2009).

However, detailed insight into how statistical structures are exploited is sparse. Specifically, despite the increased attention to this field of study (covering artificial grammar, sequence, implicit, statistical learning and predictive processing), there is no clear consensus about the neural basis of the underlying cognitive mechanisms (Conway, 2020; Dehaene et al., 2015; Denham & Winkler, 2020; Heilbron & Chait, 2018; Williams, 2020). Brain areas that have shown a significant involvement in this process are widely distributed and cover nearly the entire brain, including perceptual regions, parietal cortex, prefrontal cortex, as well as subcortical regions such as the hippocampus and basal ganglia (Conway,

2020; Henin et al., 2021; Karuza et al., 2013; Schapiro et al., 2016).

Different explanations as to why the reported activation is distributed across studies exist. One is modality (e.g., visual, auditory, tactile, motor etc.). It might be plausible that a combination of modality-specific and domain-general mechanisms interact, leading to multiple active regions across the cortex (Frost et al., 2015). Specifically, it is proposed that areas such as the prefrontal cortex or the hippocampus contribute to the domain-general mechanisms while hierarchically lower perceptual regions are the main contributors to the former. However, detailed knowledge about the exact brain regions subserving this dynamic and adaptive process is limited (Covington et al., 2018; Daikoku, 2018; Dehaene et al., 2015; Karuza et al., 2013; Meyniel et al., 2016; Schapiro et al., 2014; Strange et al., 2005; Thiessen, 2017; Williams, 2020). Consequently, this suggests that the encoding or learning of statistical regularities is not performed by one neural area but rather may be supported by multiple regions working in parallel (Chao et al., 2018; Conway, 2020; Covington et al., 2018; Frost et al., 2015; Kikuchi et al., 2017; Williams, 2020). Concerning this, it is not yet clear how this process exactly implements across areas and modalities. That is to say, whether there is one general mechanism located in a domain-general area accessible for each modality-specific system or whether each modality-specific system has its individual and optimised mechanism accessed by higher-cortical areas (Conway, 2020).

In this context, the design of the study paradigm is highly relevant because the structure of the employed stimulus sequence (for example, pure tones or linguistic material for auditory stimuli) or its complexity (ranging from deterministic to random structures) have a great effect on the engaged brain regions (Conway, 2020; Frost et al., 2015, 2019; Saffran & Kirkham, 2018; Williams, 2020). This variability makes it additionally difficult to draw generalisations from the literature. A particularly relevant case is how this mechanism operates on purely random stimuli when there is no clear meaningful relation between events. Evidence from behavioural studies suggests that we are prone to perceive patterns in random phenomena (i.e., we are prone to perform type I errors; Falk & Konold, 1997). This points towards the indication that brain mechanisms constantly operate in the attempt to make sense of the environment, regardless of whether there is any objective or behavioural relevance. Further, paradigms can either actively or passively expose subjects to stimuli. During active exposure, it is likely that attention-independent and attention-dependent systems operate in parallel. It is proposed that the implicit or attention-independent system is always active and can automatically encode adjacent or local dependencies. The explicit or attention-dependent system, on the other hand, only engages when selective attention or working memory becomes important such as for encoding or learning of non-adjacent dependencies, global patterns, cross-modal dependencies or rule-based processing (Bekinschtein et al., 2009; Conway, 2020). However, how exactly these systems are implemented is unclear. One hypothesis is that both mechanisms can operate in an antagonistic as well as in a synergistic manner (Conway, 2020). Further, previous studies have shown that during passive exposure, the brain can automatically encode simple structures that

are global without intention or awareness (Bekinschtein et al., 2009; Dürschmid et al., 2016; Williams, 2020).

Consequently, there is a need for studies to further illuminate the fundamental neural mechanisms of encoding statistical structure. Research that deploys appropriate neurophysiological methods is critical in that regard since the ratio of behavioural to neuroimaging experiments is greatly imbalanced. Especially neuroimaging studies that examine passive or implicit encoding of patterns are missing (Williams, 2020). Particularity, rendering sub-cortical regions such as the hippocampus during implicit learning tasks would be substantial since knowledge about their contribution is little, but is suspected to be fundamental (Billig et al., 2022; Covington et al., 2018; Frost et al., 2019; Schapiro et al., 2016). There might be also a necessity for novelty when it comes to paradigm designs. In previous studies, statistical parameters are often manipulated such that a high contrast between frequently and rarely occurring patterns is created (Huettel et al., 2002; Koelsch et al., 2016; Maguire et al., 2019; Mittag et al., 2016; Pelucchi et al., 2009; Saffran et al., 1996). Less structured or unstructured stimuli, on the other hand, are rarely deployed. In such stimuli, the chance of occurrence is similar or equal across event types. Identifying transient structures in such scenarios is more difficult in comparison to scenarios with highly biased probabilities. For instance, repeatedly tossing a fair coin versus tossing a highly biased one. The fair coin produces a maximally random sequence (i.e., an unstructured stimulus), whereas the sequence of the biased coin exhibits more structure since one symbol occurrence dominates over the other. Accordingly, it is more difficult to infer the probabilities of the fair coin in comparison to the biased one (Hahn & Warren, 2009; Warren et al., 2018). As addressed above, this has presumably a considerable effect on the modulated brain regions.

The need for more research with neurophysiological recordings and the need for more diverse stimuli raises the question of whether the current methodologies and approaches are proper for investigating the collected data. Are there more efficient methods to analyse recordings other than conventional approaches (stemming from statistics such as t-test, ANOVA, etc.) that primarily focus on amplitude, power, phase or latency? Alternative approaches such as neural frequency tagging (Henin et al., 2021; Picton et al., 2003) or the increased usage of artificial neural networks (Echeveste et al., 2020) or information-theoretical principles such as Shannon-entropy or mutual information (Canales-Johnson et al., 2021; Ince et al., 2017; Timme & Lapish, 2018) proved to be worthwhile in addressing fundamental questions of neuroscience. This motivates continuing the exploration of novel approaches to push for understanding the human brain as thoroughly as possible.

Specifically, a core task of the brain is to continuously encode, integrate and store *information* coming from the steady input stream of the sensory organs. Given the analogy between these mechanisms and principles from information theory (Piasini & Panzeri, 2019; Timme & Lapish, 2018), it appears plausible to employ ideas from the latter to analyze neurophysiological signals. This mathematical theory provides multivariate analysis tools, is not bound to a single type of data, is model-independent (i.e., it does not require assumptions

about the data itself) and can capture nonlinear interactions (Ince et al., 2017; Li & Vitányi, 2008; Piasini & Panzeri, 2019; Timme & Lapish, 2018). In particular, the branch of algorithmic information theory would be especially suited because it enables estimating the absolute value of the information contained in individual brain responses. While algorithmic information theory has been applied to discriminate between states of consciousness measured with neuroimaging techniques such as electroencephalography (EEG), intracranial EEG (iEEG), magnetoencephalography (MEG) or functional magnetic resonance imaging (fMRI) recordings (Canales-Johnson et al., 2020; Schartner et al., 2015, 2017; Sitt et al., 2014), its use in neuroscience is limited despite its clear potential.

1.1 Research Questions

This thesis aimed to better understand how the brain functions and to develop methods to broaden our understanding of these brain functions. Specifically, I derived two research questions under the main hypothesis that the brain continuously encodes structure within the sensory input stream and based on the relevance for further research (given the issues stated above):

Question 1 *How does the brain operate when implicitly confronted with an unstructured auditory sequence?*

Current neurophysiological evidence about how the human brain implicitly exploits statistical structure is sparse. This is especially relevant for environments showing little structure. One such scenario is environments with random features. It is plausible that various active brain regions are needed to build internal models of such complex conditions. Analysing such scenarios could constitute a contrast to more simple conditions and help to gain further insights about individual contributions of brain areas. Sensitive methods such as iEEG are a crucial component in addressing this issue because they can capture brain activity in detail. Simultaneously, by addressing the shortcomings outlined above, open questions to research fields such as the perception of randomness, statistical learning or predictive processing can be approached.

Question 2 *To what extent is algorithmic information theory an apt approach to effectively represent neurophysiological signals?*

The emergence of neuroimaging methods such as EEG, iEEG, MEG or fMRI opened up new epistemic spaces and constitute fundamental tools in modern neuroscience. As outlined, based on these tools novel issues concerning cortical information representation and processing arose. Studying these information-related issues suggests assessing the information contained within neurophysiological signals. Especially the branch of algorithmic information theory appears to be a suitable candidate to do so for it enables estimating the absolute information contained in individual brain responses

captured with neuroimaging methods. This research question examined whether algorithmic information theory with its universal applicability constitutes an apt candidate to represent signals. An evaluation of this question draws on empirical comparison to conventional methods.

This thesis is a collection of papers. While I aimed to pin down the overall thesis and presented the aim of this work in the current chapter, in the following chapter (Chapter 2), I outline relevant background information about the underlying perceptual theories and concepts. In Chapter 3, I then depict the methods used and developed in the thesis. This is followed by a summary of the resulting research (Chapter 4) and a discussion (Chapter 5) of the outcomes, where I evaluate the findings in terms of the research questions. Lastly, I present a conclusion (Chapter 6) and all research articles that resulted from this project.

Chapter 2

Background

In this chapter, I outlined the core concepts of the thesis: First, the two treated research fields (predictive processing and statistical learning) anchored in perception and learning are introduced. This is followed by a general perspective on how the brain possibly encodes patterns in the environment. Lastly, considerations in terms of randomness and its perception are put forward, followed by a brief introduction to information theory.

2.1 Predictive Processing

Doubtlessly, predictive processing or predictive coding is the current dominant theoretical framework for the study of perception, action and learning (Fig. 2.1, Bastos et al., 2020; Denham & Winkler, 2020; Friston, 2018). Perception deals with the processing of information from external or internal input conveyed by our senses. As a first general theory, predictive processing aims to explain both psychological and neural aspects of perception, that is, accounting for both behavioural and neurophysiological processes (Denham & Winkler, 2020). According to this theory, human brains perceive through internal models of the environment by constantly making predictions based on prior expectations in order to explain away an incoming sensory signal. The better the fit between top-down predictions and lower-level sensory input, the higher the probability that the prediction accurately infers the hidden causes of the lower-level activity. Hence, the brain aims to minimise the discrepancy, that is prediction error or surprise between prediction and sensory input. This occurs within a hierarchical bidirectional cascade of cortical processing ranging from an interaction between individual neurons up to communication between large populations of neurons (i.e., brain areas). Based on evidence from the frequency decomposition of EEG and iEEG from humans and other animals, it is hypothesized that top-down alpha (8 to 14 Hz) or beta (15 to 30 Hz) activity carries descending information and regulates the processing of bottom-up inputs, whereas ascending information is served by the gamma band (>30 Hz). In this framework, lower frequencies carry slower top-down predictions, inhibiting faster gamma activity that processes external inputs (Arnal & Giraud, 2012; Bastos et al., 2012, 2015, 2020; Chao et al., 2018; Heilbron & Chait, 2018; Wacongne et al., 2011). Accordingly, predictions trying to infer a complex cause should be processed at higher cortical levels (e.g., the pre-frontal cortex) while simpler rules should be implemented in lower cortical areas (e.g., within the subcortical or primary auditory pathway). While predictive processing theories afforded novel explanations on how brain networks process and integrate information based on empirical investigation, its extensive proof still lies ahead (Denham & Winkler, 2020; Friston, 2018; Heilbron

& Chait, 2018). Especially how the neuro-architectonic structure implements such a mechanism in the brain is not clear yet. Specifically, while there are several studies that index prediction errors within neurophysiological recordings, the demonstration of predictions is sparse (Arnal & Giraud, 2012; Bastos et al., 2012; Carbajal & Malmierca, 2018; Clark, 2013; Denham & Winkler, 2020; Friston, 2010; Heilbron & Chait, 2018; Parras et al., 2017; Rubin et al., 2016).

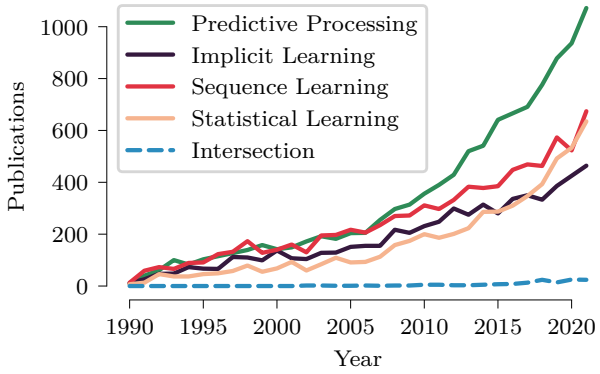


Figure 2.1: Number of yearly publications from 1990 to 2021 containing the keywords in the legend within psychology and neuroscience extracted from the data bank of Web of Science. The dashed line represents the yearly articles containing both predictive processing and statistical learning as keywords.

2.2 Statistical Learning

Ultimately, the field of statistical learning arose from a study examining language acquisition (Saffran et al., 1996). This study revealed that eight-month-old infants are sensitive to conditional regularities between adjacent speech sounds after only two minutes of exposure to a fake language. This evolved into a rapidly growing research field (Fig. 2.1) that covers "all phenomena related to perceiving and learning any forms of patterning in the environment that are either spatial or temporal in nature" (Frost et al., 2019). Accordingly, it considers the brain's capability to extract regularities between features and objects that co-occur in the environment over space and time (Conway, 2020; Frost et al., 2015; Saffran, 2020; Thiessen, 2017; Turk-Browne, 2012; Williams, 2020). In fluent speech, for instance, conditional regularities between syllables constitute patterns that help to identify word-like units, generating candidate words available for mapping to meaning (Arnal & Giraud, 2012; Conway, 2020; Dehaene et al., 2015; Koelsch et al., 2016; Saffran, 2020; Thompson & Newport, 2007). Hence, monitoring conditional statistics enables higher-order learning processes to make sense of said. Based on that, it is even hypothesised whether that could be a predecessor component upon which human language evolved (Kikuchi et al., 2017; Saffran

et al., 1996). Consequently, statistical learning constitutes a major theoretical construct in cognitive science that is not limited to speech perception or language acquisition but is also applicable to studies in vision, audition, motor planning, event processing, reading, semantic memory or social cognition (Frost et al., 2019; Saffran & Kirkham, 2018).

As mentioned in the introduction, this patterning is indispensable for the internal perceptual models of the environment, which again is a core concept of predictive processing. Accordingly, while emerging from different origins, predictive processing and statistical learning evolved into grand theories that are closely related. Other related fields in this regard are artificial grammar, sequence or implicit learning (Fig. 2.1). Implicit learning is of special interest to this thesis since it considers learning without any intention or awareness involved, which is also the focus of this project.

From these relations follows that both predictive processing and statistical learning face similar problems when it comes to pinning down the neurophysiological mechanisms. However, studies orientating simultaneously at both fields are sparse (Fig. 2.1). While for predictive processing, there is a significant amount of studies that employ neuroimaging techniques (however not so many that investigate statistics as perceptual variables), behavioural studies outweigh the field of statistical learning. Accordingly, in statistical learning, the brain's capability to effectively identify regularities in ongoing sensory input has been sufficiently observed on a behavioural level in advanced experimental setups across multiple domains, yet the neurological understanding is less clear. Specifically, the link to the underlying mechanisms or its motivation through which these patterns are detected in the first instance (Barascud et al., 2016; Conway, 2020; Covington et al., 2018; Daltrozzo & Conway, 2014; Saffran, 2020; Schapiro et al., 2014; Thiessen, 2017).

2.3 Perceptual Variables

The brain is sensitive to two types of regularities, which are either distributional or conditional (Saffran, 2020; Thiessen, 2017). Accordingly, they describe the absolute frequency of individual event types or the relationship across multiple events. The most prominent statistics of the former category are item frequency (IF) or alternation frequency (AF). Several studies have demonstrated that the brain is sensitive to both regularities, depending on the structure or dynamics of the stimuli (Maheu et al., 2019; Saffran, 2020; Thiessen, 2017). Given a discrete binary sequence (e.g., XYXYXYXXYXYXYYXYYXXYXYX) IF describes how frequently each event type occurs relative to the total amount of events within the sequence, i.e., it can be defined as

$$\mathrm{IF(X)} = \mathrm{p(X)} = \frac{\mathrm{n(X)}}{\mathrm{n(all)}} \tag{2.1}$$

for the event X and the frequency operator n. Consequently, in the above sequence, the IF for the event X becomes IF(X)=1−IF(Y)=12/23=0.52. IF,

however, neglects how the individual event types are distributed within the sequence. AF, on the other hand, looks at the number of adjacent transitions between events. For a binary sequence, it is defined as

$$\mathrm{AF} = \mathrm{p(alternation)} = \frac{\mathrm{n(X \to Y) + n(Y \to X)}}{\mathrm{n(all) - 1}}. \tag{2.2}$$

Accordingly, AF models the number of alternations relative to the total number of possible alternations and becomes p(alternation) = 1−p(repetition) = 18/23=0.78 for the sequence above. As the name suggests, both probabilities model the distribution of a sequence. However, while IF is not accounting for the order, AF does not cover directionality.

A perceptual variable considering both is the transitional probability (TP). As a conditional regularity, it entails information regarding AF, IF and directionality. In general, it can be defined as the ratio of the directional co-occurrence of events given their frequency (Maheu et al., 2019; Pelucchi et al., 2009; Thiessen, 2017; Thompson & Newport, 2007). For each possible transition, it is defined as

$$\mathrm{TP(Y, X) = p(Y|X) = \frac{n(X \to Y)}{n(X)}}\,. \tag{2.3}$$

In the case of a binary sequence, this results in a TP or stochastic matrix of size $\mathbb{R}^{4\times4}$, i.e.,

$$\mathrm{P} = \begin{pmatrix} \mathrm{p(X|X)} & \mathrm{p(X|Y)} \\ \mathrm{p(Y|X)} & \mathrm{p(Y|Y)} \end{pmatrix}. \tag{2.4}$$

Therefore, TPs describe how likely one event predicts another. For the example sequence above, this results in p(X|Y)=1−p(Y|Y)=9/11=0.82 and p(Y|X)=1−p(X|X)=9/11=0.82. When estimating the TPs along a sequence in an incremental fashion, that is to say for each sample along the time axis, usually, the principle of indifference is applied for the initial value, leading to equal probabilities for all initial transitions. See Section 3.2 for more information on the framework of TP estimation in this project.

Sensitivity to conditional regularity between events has been observed in humans (Domenech & Dreher, 2010; Henin et al., 2021; Higashi et al., 2017; Koelsch et al., 2016; Leonard et al., 2015; Maguire et al., 2019; Maheu et al., 2022; Meyniel & Dehaene, 2017; Meyniel et al., 2015, 2016; Mittag et al., 2016) and animals (Lu & Vicario, 2014; Yaron et al., 2012). Because events in the environment rarely occur independently, this pattern extraction is necessary for the fast and efficient processing of sensory information. In past research, it was suggested that this regularity is a core structure for implicitly encoding patterns and most of the statistical learning research focused on designing sequences by manipulating TPs (leading to high and low TP patterns) (Frost et al., 2019; Thiessen, 2017).

2.4 Perception of Randomness

In the current thesis, randomness and its perception play an important role. A key question in cognitive science is whether the brain can be aware of randomness or if it always searches for patterns. In other words, does it always try to make sense of or exploit a given sequence? That would follow the notion that the brain continually maintains a detailed representation of the structure of the unfolding sensory input and that this representation shapes the processing of incoming information (Clark, 2013; Denham & Winkler, 2020; Friston, 2008; Heilbron & Chait, 2018; Southwell & Chait, 2018).

Applying this notion to randomly distributed sequences, it becomes plausible that to make predictions about following events the brain exploits regularities within the random sequences. By doing so, the brain performs a type I error or perceives patterns in random phenomena (Falk & Konold, 1997). An example demonstrating this behaviour is depicted in Fig. 2.2. Each graph possesses 100 points. From intuition follows that the right graph is random, whereas the left plot exhibits too many repetitions or clusters of points to be random. These clusters enable the identification of patterns or even abstract forms. In reality, however, the dots on the left-hand side are randomly distributed. The right one distributes the points uniformly by dividing the plot into 100 boxes where each box contains one point, and each location is random in that box. This tendency or intuition to perceive patterns in random data is termed clustering illusion. The term illusion is chosen because the effect does not vanish by repeated examination, which is equal to perceptual illusions. Further, random distributions seem to have too many clusters of the same consecutive outcome, such that humans struggle to accept their true origin (Gilovich, 1991).

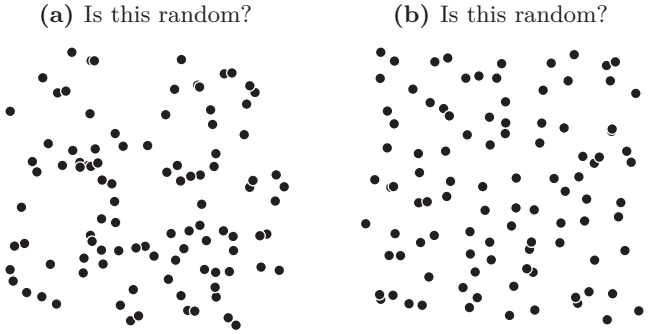


Figure 2.2: Two graphs with 100 points each. **a**: Randomly distributed points. **b**: The points obey a uniform distribution by dividing the plane into a grid of 100 boxes. One point is randomly located in each box.

The very nature of randomness enables the identification of patterns, especially in large data sets. That is consistent with the Ramsey Theory, which postulates that every large set of numbers exhibits highly regular patterns

(Graham & Spencer, 1990). Other terms addressing the same tendency as clustering illusion are apophenia or pareidolia, where the former is defined by Conrad (1958) as an "unmotivated seeing of connections [accompanied by] a specific feeling of abnormal meaningfulness". This phenomenon allows us to see shapes in clouds, identify words in noise or perceive a face on Mars (Nasa, 2001).

Committing a type I error is also plausible from an evolutionary point of view. Concerning survival, it is more beneficial to interpret a stream of wind waving through a dense hedge as something dangerous than to interpret it the other way around. That is to say, performing a type II error by believing it is only the wind or assuming something is random when it is not. Another example is the hot hand in basketball or gambler's fallacy (Falk & Konold, 1997; Gilovich et al., 1985; Kahneman & Tversky, 1972; Warren et al., 2018). Repetition is recognised as something non-random, such that the performance of a basketball player is ascribed as "hot" or "on fire" when one has a streak of hitting shots (Gilovich, 1991; Gilovich et al., 1985). Players, coaches and fans believe that once a player makes a basket their chances of hitting the next one increases. The same cognitive bias applies to gamblers who tend to assume that after a long sequence of, for instance, black in roulette, it becomes more likely that the outcome of the next spin is red. Yet, in each trial, both outcomes are equally likely. Assuming that the local and global sequences should share the same properties is known as the gambler's fallacy (Kahneman & Tversky, 1972; Warren et al., 2018). The future is not influenced by the past in these cases.

Another factor fostering the encoding of patterns in random stimuli is the length of a perceived sequence. Here, the terms insensitivity to sample size or law of small numbers comes into play. That describes the human tendency to conclude about the whole population given a small sample size (Falk & Konold, 1997; Kahneman & Tversky, 1972; Zhao et al., 2014). For example, for a coin toss, the probability for head and tail is p(head)=p(tail)=0.5 for each trial. Yet, (longer) repetitions of one symbol are assessed as non-random without considering that this behaviour could change with increasing sample size. The opposite would be over-alternations. In that case, the two states alter consecutively for a long period, and for both the generation and perception of sequences this is considered to be maximally random (Falk & Konold, 1997; Kahneman & Tversky, 1972). In the example of Fig. 2.2, such an over-alternation can be the distance between dots which is too uniformly distributed in the right graph to be random. Needless to say, there is also the probability to exactly obtain this plot by distributing one hundred points randomly, but this again addresses the question of how randomness is precisely defined and how to measure it, which goes beyond the scope of this thesis.

From these examples of how we perceive randomness as a meaningful pattern follows that the brain is not aware that the underlying structure is random or meaningless. Hence, in scenarios where there is no attention involved, the brain should not be able to implicitly identify and subsequently disregard something as being of random nature, to begin with. Accordingly, patterns should always be implicitly exploited to maintain a detailed representation of the structure of the unfolding sensory input and then made use of by the brain. A study supporting

this claim is Southwell and Chait (2018). In this study, the difference between recorded brain responses to sounds that deviate from standard tones following either regular or random ordered standard tones was examined. Southwell and Chait concluded that responses were substantially larger following regular sequences relative to random ones. Hence, the degree of predictability in the ongoing sequence context plays a role in how a deviant event is treated.

2.5 Information Content

The second research question considers to what extent information theory qualifies to treat neuroimaging recordings through an empirical approach. The choice as to why to use information theory is apparent. On an abstract level, neural systems constantly encode, integrate, and store information originating from the steady input stream of the sensory organs. Information theory has been widely applied to neuroscience research as it offers multivariate analysis tools, is not bound to a single type of data, is model-independent (i.e., does not require assumptions about the data itself) and can capture nonlinear interactions (Ince et al., 2017; Piasini & Panzeri, 2019; Timme & Lapish, 2018). Specifically, principles of Algorithmic Information Theory or more precisely, Algorithmic Complexity or Kolmogorov Complexity (K-complexity) have been employed to discriminate between states of consciousness measured with EEG, iEEG, MEG or fMRI recordings (Canales-Johnson et al., 2020; Schartner et al., 2015, 2017; Sitt et al., 2014; Varley et al., 2020). However, its use is less popular than the classic information theory after Shannon (1948). As opposed to the latter, K-complexity allows estimating the information content or complexity of an individual object. To give a short overview of both theories, firstly, ideas of the K-complexity approach are outlined, followed by the sketch of information content definition after C.E. Shannon's theory of communication.

2.5.1 Algorithmic Complexity

Algorithmic information theory was independently developed by R.J. Solomonoff (Solomonoff, 1964), A.N. Kolmogorov (Kolmogorov, 1968) and G. Chaitin (Chaitin, 1969). Regarding this theory, the information content of an individual object can be estimated by measuring the length of its shortest description (Grünwald & Vitányi, 2004; Li & Vitányi, 2008). For example, the sequence

ABABABABABABABABABABABABAB

can be described as "AB repeats 27 times". The structure of the sequence

AABBABBBBAABBAAABAAAABBBAA

on the other hand, is less clear. Intuitively, there appears to be no simple description such that the second sequence is more complex or exhibits a higher information content than the first sequence. More formally, K-complexity is the ultimate compressed version or minimum description length of an object or its

absolute information content (Li & Vitányi, 2008). If the minimum description length is short (long), an object is characterised as "simple" ("complex"). As evident from the two example sequences, there can be scenarios where it is rather difficult to infer a rule that describes an object with minimum description length. In general, it is not possible to compute the theoretically ideal K-complexity. Therefore, when applying this theory, an estimate needs to be chosen, yielding an upper-bound approximation (an estimate of the description length is always longer than the ultimate theoretically possible minimum description length). Possible estimation approaches are conventional lossless data compression programs such as gzip, bzip2 or the Lempel–Ziv–Markov chain algorithm (LZMA) (Li & Vitányi, 2008; Li et al., 2004). See Fig. 2.3 for example signals and their compressed value estimated with gzip.

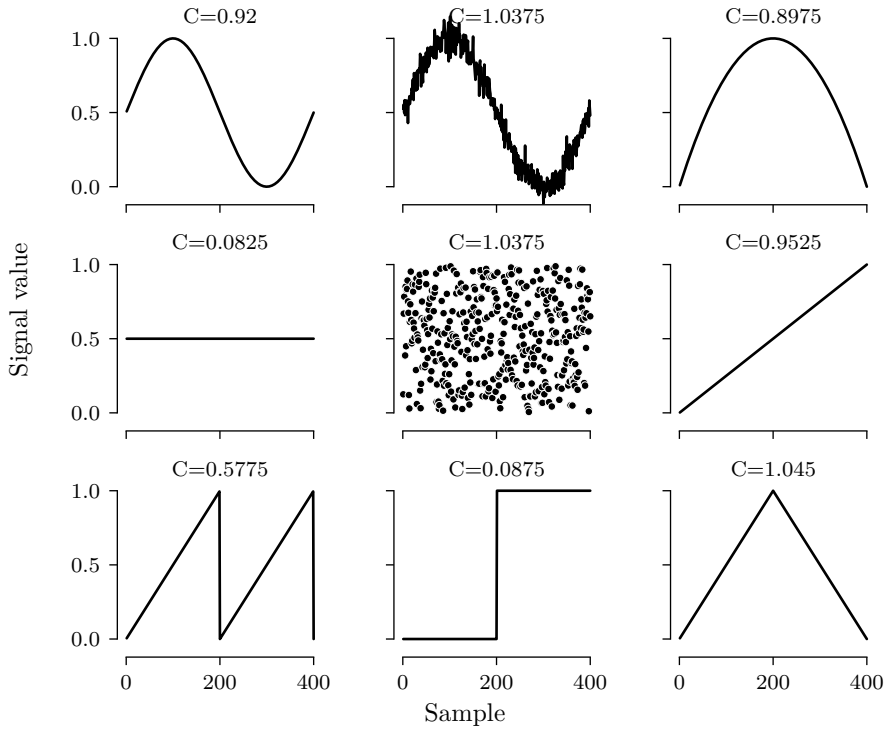


Figure 2.3: Example functions or signals that underwent compression with gzip. The two signals of column two in rows one and two resulted from adding Gaussian white noise to the respective signals in column one. Note how only regularities such as the saw tooth function enable the compressor to exploit structures along the sample axis. The triangular pulse, conversely, does not allow any compression. The signal's complexity value (C) was computed by dividing its compressed version by its original signal size. For complexity values greater than one, the compressor cannot reduce the size of the respective signal.

2.5.2 Encoded Information

In neuroscientific studies, it is common to have two or more conditions (for example, control and test conditions) that need to be compared with each other. The K-measure so far yields an estimate of the absolute information content of a recording. Fortunately, based on the K-complexity, various metrics were derived that allow a pairwise comparison between two objects. One instance is the Normalized Information Distance or its estimation counterpart, the Normalised Compression Distance (NCD). The NCD allows for the comparison of different pairs of objects with each other and suggests similarity based on their dominating features (or a mixture of sub-features) (Cilibrasi & Vitányi, 2005; Li & Vitányi, 2008; Li et al., 2004). This measure has been applied to, for example, cluster analysis, virology (even to analyse the SARS-CoV-2 virus by two of the original authors behind this measure (Cilibrasi & Vitányi, 2022)), language, music and literature (Li & Vitányi, 2008), but as far as know has only been applied once in the field of neuroscience (i.e., for brain-computer interfaces, Sarasa et al., 2019). For a pair of strings (x, y), the $\mathrm{NCD}(x, y)$ is defined as

$$\mathrm{NCD}(x, y) = \frac{C(xy) - \min(C(x), C(y))}{\max(C(x), C(y))}\,, \tag{2.5}$$

with $C(xy)$ denoting the compressed size of the concatenation of x and y, and $C(x)$ and $C(y)$ their respective size after compression (Cilibrasi, 2007; Cilibrasi & Vitányi, 2005; Li & Vitányi, 2008; Li et al., 2004; Vitányi et al., 2009). Further, the NCD is non-negative, that is, it is $0 \leq \mathrm{NCD}(x, y) \leq 1 + \epsilon$, where the ϵ accounts for the imperfection of the employed compression technique. Smaller NCD values suggest similar objects. Higher values suggest rather different objects. There is an intuitive interpretation of NCD (Cilibrasi, 2007; Vitányi et al., 2009). For instance, given $C(y) \geq C(x)$ the compression distance becomes

$$\mathrm{NCD}(x, y) = \frac{C(xy) - C(x)}{C(y)}\,. \tag{2.6}$$

Accordingly, it becomes the ratio between the improvement of compressing y using x as a previously compressed "database" relative to compressing y from scratch (Cilibrasi, 2007; Li & Vitányi, 2008). Using the definition of the conditional compressed information $C(y|x) = C(xy) - C(x)$, it can be also written as

$$\mathrm{NCD}(x, y) = \frac{C(y|x)}{C(y)}\,, \tag{2.7}$$

that is, it is the ratio of the information x about y to the information in y (Cilibrasi, 2007).

Labelling its implementation for neurophysiological data *encoded information*, we utilised the complexity-based NCD measure to quantify information-content-based differences in neurophysiological recordings. Paper II and Paper III evaluate this approach in detail. Overall, the measure was applied to three

different neurophysiological (Paper I, Paper II and Paper III) and synthetic data sets (Paper II). We quantified differences across single trials (evoked responses) or conditions (mean response across trials) based on five different cortical activity types for each data set. The signals considered ranged from 0.35 to 2 s. See the next chapter (Section 3.3 and Section 3.3.1) for more details on the composition of the different neurophysiological signals.

The general procedure to obtain a compression version $C(x)$ of a signal x was to first represent the continuous signal by grouping its values into discrete steps (bins). Accordingly, the signal was simplified by reducing its resolution along the y-axis (Fig. 2.4; for example into 127 bins). The bins covered equal distances and were in a range between the global extrema of all recorded signals (e.g., the global extrema of a channel). After simplifying the signal, a compressor then received the indices of the bins that contained the elements of the signals (Canales-Johnson et al., 2020; Sitt et al., 2014). That is to say, each sample is mapped to a value of the binning vector. This index vector underwent then compression through a compression routine based on Python's standard library and gzip. At this point, it is worth pointing out that the length of the index vector is invariant to the number of bins. This is because the length of the signal is not changed during the binning procedure, but only the resolution along the y-axis. Consequently, each value within the index vector represents a signal sample in ascending order. Thus, the input length of the compressors remains constant.

For example, consider the signal x consisting of nine random values

$$x = \begin{pmatrix} 3.47 & 2.14 & 2.55 & -0.18 & 2.85 & 1.05 & 1.20 & 2.94 & 1.59 \end{pmatrix}^{\mathsf{T}}.$$

Using 12 bins that cover equal distances and in a range between the global extrema, the following (ascending) edges result:

$$edges = \begin{pmatrix} -\infty & 0.12 & 0.43 & 0.73 & 1.04 & 1.34 & 1.65 & 1.95 & 2.26 & 2.56 & 2.87 & 3.17 & \infty \end{pmatrix}^{\mathsf{T}},$$

Note that on each end, ∞ is added to account for machine precision. Assigning each value of the signal to the closest edge then leads to the vector

$$x_{\text{bin}} = \begin{pmatrix} 12 & 8 & 9 & 1 & 10 & 5 & 5 & 11 & 6 \end{pmatrix}^{\mathsf{T}}.$$

Accordingly, the vector x_{bin} maps each element of x to a value of the binning vector *edges*. This simplified and binned vector x_{bin} is subsequently received by the employed compressor, yielding its compressed version $C(x)$.

Across the three data sets considered (quantifying differences of two conditions for each of the five cortical activity types, see Paper III), the median encoded information was 0.7414 ± 0.0528, the minimum 0.5344 ± 0.1206 and the maximum was 0.9720 ± 0.0708.

2.5.3 Shannon Entropy

In common with algorithmic information theory, Shannon's information theory suggests ways to measure the information of an object. However, whereas for

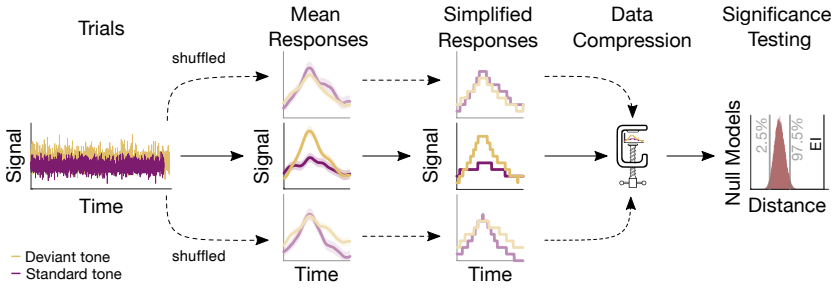


Figure 2.4: Sketch of the procedure. Based on all recorded trials a mean response for each condition is computed. The trials (i.e., not the samples within a trial) are then shuffled, resulting in surrogate mean responses. Subsequently, these signals undergo a simplification procedure, followed by compression. The output of the compression routine is the encoded information, quantifying the similarity between responses. The resulting values are then evaluated, leading to a null model distribution. This distribution serves to assess the significance of the true encoded information value.

the K-complexity, individual objects themselves are considered, the approach after Shannon ignores the meaning of the variable and is only interested in the transmission of the variable from a sender to a receiver (Grünwald & Vitányi, 2004; Li & Vitányi, 2008). That is to say, it enables estimating the uncertainty or variability of a random variable by introducing a concept called entropy or information content. This can be best understood through a random variable underlying a distribution P. A value x with low occurrence probability P(x) means it is less likely to occur. Consequently, a receiver or observer would be surprised to receive it. A value with high probability P(x), conversely, would be less surprising to the observer (Ince et al., 2017). Accordingly, the surprise for value x drawn from the distribution P(x) can be defined as $-\log_2 \mathrm{P(x)}$ (i.e., for $\mathrm{P(x)} \to 0$, $-\log_2 \mathrm{P(x)} \to \infty$ and $-\log_2 1 = 0$). The entropy is then defined as the expected (average) surprise over the distribution, that is for a discrete distribution, it is defined as

$$H(x) = -\sum_{k=1}^{\mathrm{N}} \mathrm{p}(x_\mathrm{k}) \log \mathrm{p}(x_\mathrm{k}) \,, \tag{2.8}$$

with $\mathrm{p}(x_\mathrm{k})$ being the occurrence probability for each element $x_\mathrm{k}, \ldots, x_\mathrm{N}$ of x. Consequently, if an observer draws several samples from a distribution with lower entropy means that each draw will be of little surprise for the observer because the distribution shows little uncertainty. High entropy, on the other hand, means that each sample has a similar probability, i.e., the distribution is spread out, such that the outcome of each draw is uncertain. According to this, for a discrete random variable, the maximal possible entropy exists for a uniform distribution exhibiting equal probabilities for each sample (for example tossing a fair coin compared to a highly biased one, Fig. 3.4a&3.4b). To compute the entropy of a

signal, a probability distribution needs to be estimated. One commonly used method is to group their values into bins, which is the same procedure described above in section 2.5.2. Based on the bin distribution, maximum likelihood estimation is then applied (that is, evaluating in the histogram how many times each bin occurs relative to the total amount of samples within a signal), yielding the probability distribution estimate.

Similar to the algorithmic approach, there exists an entropy measure which allows quantifying the relationship between two variables in terms of differences in entropies. Named mutual information (MI), it can be defined between two discrete random variables (x, y) with N or M outcomes as

$$\begin{aligned} \mathrm{MI}(x;y) &= H(x) - H(x|y) \\ &= H(y) - H(y|x) \\ &= \sum_{\mathrm{k}=1}^{\mathrm{N}} \sum_{\mathrm{j}=1}^{\mathrm{M}} p(x_{\mathrm{k}}, y_{\mathrm{j}}) \log \frac{p(x_{\mathrm{k}}, y_{\mathrm{j}})}{p(x_{\mathrm{k}})p(y_{\mathrm{j}})} \,, \end{aligned} \tag{2.9}$$

with the joint probability $p(x_{\mathrm{k}}, y_{\mathrm{j}})$ and the marginal probabilities $p(x_{\mathrm{k}})$ and $p(y_{\mathrm{j}})$. Important to note is that MI makes no assumptions on the distribution of the variables or the nature of the relationship between them (Ince et al., 2017; Timme & Lapish, 2018).

Chapter 3

Methods

In this chapter, an overview of the methodologies is presented, which I used for investigating the research questions, and outlines the main experiment of this thesis. Individual methodologies for each article are presented in Chapter 4, where each respective publication specifies the details.

3.1 Implicit Listening Paradigm

At the core of this thesis lies a multi-dimensional oddball paradigm named "Optimum-1 Paradigm" (Näätänen et al., 2004). This paradigm constitutes an ideal candidate for investigating the outlined research questions. Designed as an unattended listening task, it targets the auditory system, which is a great testbed for examining how the sensory input is processed, availing on the auditory Mismatch Negativity (MMN). This again is an index of automatic information processing occurring in the auditory cortices and assumedly to be the most well-studied neural signature of surprise (Carbajal & Malmierca, 2018; Denham & Winkler, 2020; Heilbron & Chait, 2018; Paavilainen, 2013).

The oddball paradigm used in this study consists of one standard tone and five tones deviating from the standard tone (Fig. 3.1). Standards had a duration of 75 ms with 7 ms up and down ramps and consisted of three sinusoidal partials of 500, 1000 and 1500 Hz. Deviants varied relative to the standard in the perceived sound-source location (left or right), intensity (±6 dB), frequency (550, 1100 and 1650 Hz or 450, 900 and 1350 Hz), gap (25 ms silence in the middle) or by a shortened duration (1/3 or 25 ms shorter). As depicted in Figure 3.1, three of the five deviant types have an individual variation, namely location, intensity and frequency deviants.

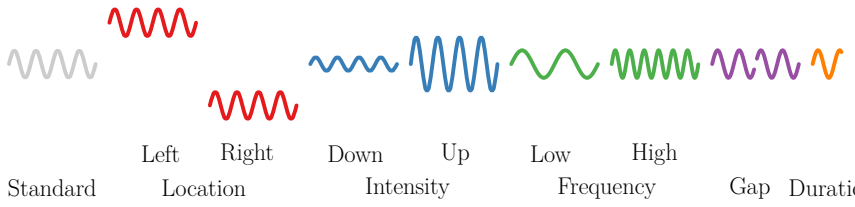


Figure 3.1: Illustration of the different tone types. Aside from a standard tone, five deviant types are presented. Out of the five deviant sounds, three possessed individual variations.

During the sequence presentation, each standard is followed by a deviant tone (Fig. 3.2). As each deviant type only varies with one property relative to the standard, the memory trace of the standard tone is strengthened, with the

result that each deviant tone presentation elicits an MMN or surprise response (Näätänen et al., 2004). Within a set of five consecutive deviants (that is, ten tones in total) each of the five types were presented once. In consecutive sets, the same deviant type did not repeat from the end of one set to the beginning of another. For the three deviants that had two stimuli versions, each version occurred equally often (p=0.5). Except for deviants varying in duration, all tones had a duration of 75 ms and were presented every 500 ms in blocks of 5 min consisting of 300 standards and 300 deviants tones. Lastly, at the beginning of each block, 15 standard tones were played.

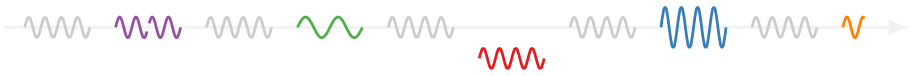


Figure 3.2: Snippet of the multi-dimensional auditory oddball paradigm stream.

As the main aim of this project was to examine the brain's capability to capture automatic, stimulus-driven processes, participants were asked not to pay attention to the sounds while reading a book or magazine. They completed three to ten blocks, providing at least 1800 trials (i.e., evoked event responses).

The sequence exhibits equal probabilities for the different deviant types. Consequently, based on that random aspect within the paradigm, clusters of repeating deviant types and patterns between deviant types naturally emerge. These patterns arise because a finite stream naturally entails temporal patterns of event occurrence. This is in contrast to a stream where the length converges towards infinity (for example, a finite coin toss always shows patterns, but when observed over an infinite horizon no patterns exist).

Based on the main hypothesis that the brain always encodes patterns, the brain should implicitly encode the naturally arising temporal structures to build an internal model of the sensory input (Fig. 3.3).

3.2 TP estimation

To systemically investigate this statement I hypothesized in this thesis that the brain shows sensitivity to TPs mirroring the basic structure of statistical patterns during the stream of sounds. In the analysis, the theoretically ideal TPs were incrementally estimated between adjacent deviants in the fashion of an ideal observer (Eq. 2.3). At each given deviant event (trial), TP estimates were updated based on all previously presented deviant stimuli. Consequently, TPs dynamically evolved along the course of the experiment (Maheu et al., 2019). That can be best imagined by considering a coin toss. As depicted in Fig. 3.4, there is always a variance from toss to toss for both a fair coin (Fig. 3.4a) or a biased coin (Fig. 3.4b). The same can be observed for the employed stimuli (Fig. 3.4c).

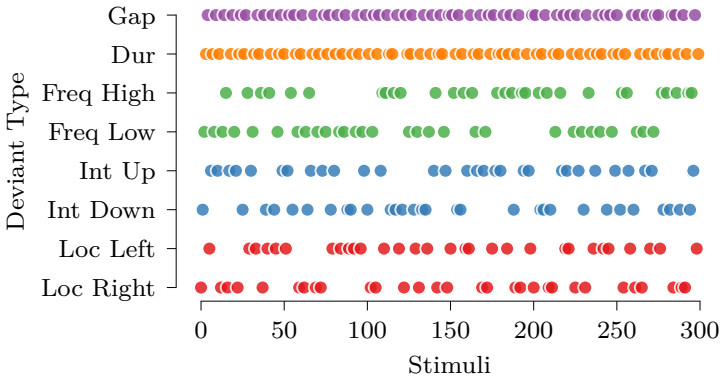


Figure 3.3: Distribution of deviant tones over the sequence. Random patterns arise by transitioning from one deviant type to another. Gap and duration deviants only have one version, while two variations for the remaining deviant tone types exist.

3.3 iEEG

The neurophysiological activity in focus for this project was iEEG. While iEEG recordings with their low signal-to-noise ratio (SNR) give access to a unique temporal and spectral resolution, the number and placement of electrodes are exclusively driven by clinical requirements. Consequently, to make statements not locally constrained to a specific region, there must be enough subjects to obtain distributed electrodes across the cortex. For the paradigm introduced above, iEEG was recorded from 22 subjects (15 at El Cruce Hospital and seven at Oslo University Hospital) that were self-reported normal hearing adults with drug-resistant epilepsy (mean age 31 years, range 19 to 50 years, 6 female). They were potential candidates for resective surgery and as part of their pre-surgical evaluation underwent invasive intracranial electrocorticography (ECoG) or stereoelectroencephalography (SEEG). Intracranial electrodes were temporarily implanted to localise the epileptogenic zone and eloquent cortex. The electrode shafts were implanted in different regions of the frontal, temporal, occipital, central parietal and subcortical structures (for further details see Blenkmann et al., 2019; Fuhrer, Glette, Ivanovic, et al., submitted).

To process the raw iEEG, the data was manually inspected. Electrodes or epochs with epileptiform activity or abnormal signals were then removed. Electrodes located in lesional tissue or tissue that was later resected were excluded. Subsequently, bipolar channels were computed as the difference between signals recorded from pairs of neighbouring electrodes in the same electrode array. The resulting data was then low-pass filtered at 180 Hz. Line noise was removed using bandstop filters at 50, 100 and 150 Hz. Data was then segmented into 2000 ms epochs (−750 ms before and 1250 ms after tone onset) and demeaned.

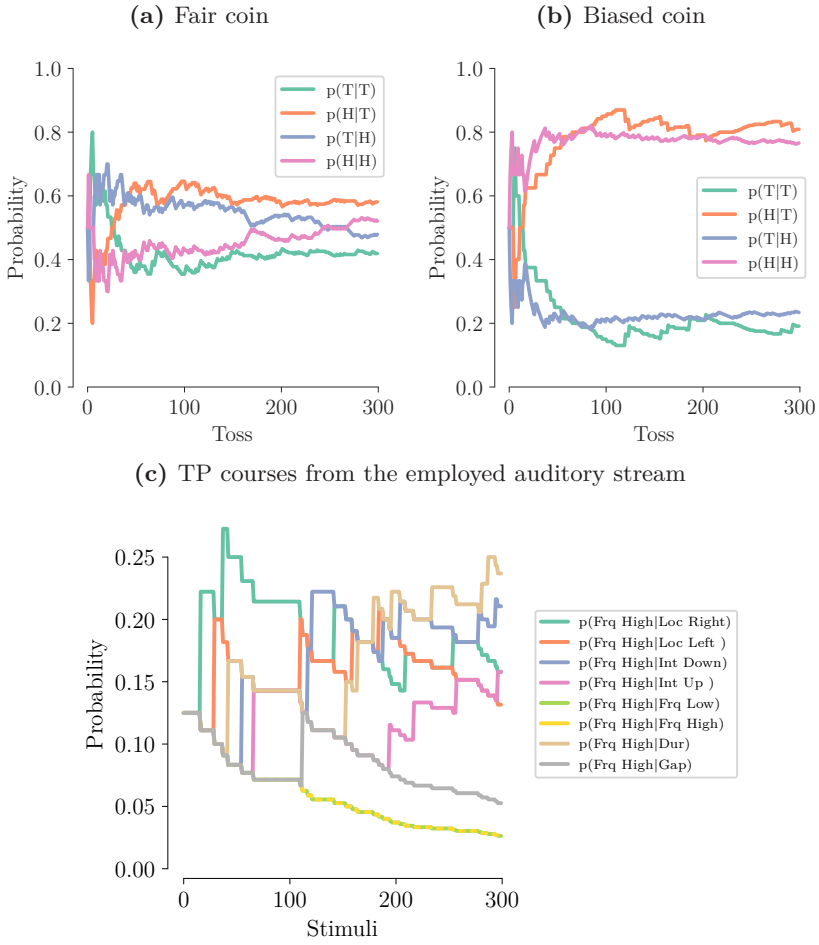


Figure 3.4: Example TP courses, where the principle of indifference is applied at the beginning. **a**: The TP courses emerge from a coin toss with equal probabilities (i.e., p(H)=p(T)=0.5). **b**: TP courses emerge from a coin toss with biased probabilities (i.e., p(H)=0.8, p(T)=0.2). **c**: Example TP courses resulting from the incremental estimation across the stimuli. The graph represents the TPs across the sequence for the "frequency high" deviant type. Note that the same deviant cannot occur adjacently (green and yellow line). Further, all occurring TPs should converge to $1/6$=0.167.

Based on the segmented recordings, cortical activity types of theta (5 to 7 Hz), alpha (8 to 12 Hz), beta (12 to 24 Hz), high-frequency (HFA; 75 to 145 Hz) and broadband event-related potential (ERP) were extracted. The procedure was the same for all data sets considered (see Paper III for numerical details of each data set).

To obtain the ERP, data was low-pass filtered at 30 Hz using a sixth-order Butterworth filter. Theta, alpha and beta frequency bands were extracted from the demeaned signals using wavelet time-frequency transformation (Morlet wavelets) based on convolution in the time domain (Oostenveld et al., 2011). All trials were then baseline corrected by subtracting the mean amplitude of the baseline period of each trial and frequency band from the entire trial (the paradigm above had a baseline window of −100 to 0 ms relative to tone onset).

3.3.1 High-frequency Activity

To obtain the HFA of the recordings, we filtered the preprocessed data into eight bands of 10 Hz ranging from 75 to 145 Hz by use of bandpass filters with Fieldtrip (Oostenveld et al., 2011). Next, the instantaneous amplitude signal of each filtered signal was computed by applying a Hilbert transform to the filtered time series leading to the analytic signal (Foster et al., 2016)

$$\zeta(t) = X(t) + iH(X(t))\,. \tag{3.1}$$

The analytic signal $\zeta(t)$ for each band is a complex-valued time series. The real component $X(t)$ is the original band passed signal, and the imaginary component $iH(X(t))$ is the Hilbert transform of $X(t)$, defined as

$$H(X(t)) = \frac{1}{\pi}\text{P.V.}\int_{-\infty}^{+\infty}\frac{X(t)}{t-\tau}d\tau\,, \tag{3.2}$$

using the Cauchy principal value denoted as P.V. (Canales-Johnson et al., 2020; Foster et al., 2016). The Cauchy principal value allows the assignment of values to improper integrals. This can occur because of the discontinuous convolution term $1/\pi t$ which is undefined at zero and thus possibly renders the integral improper. We then obtained the analytic amplitude time series $a(t)$ or signal envelope corresponding to specific frequency bands by applying Pythagoras' Theorem, that is to say

$$a(t) = \sqrt{X(t)^2 + H(X(t))^2}\,. \tag{3.3}$$

To obtain one time series across all eight frequency bands we calculated their mean amplitude value. As the last step, the respective time series were normalised by dividing them by a mean baseline period which was computed from all trial recordings. This resulted in a normalised measure which was relative to the baseline activity and termed HFA. HFA reliably represents the underlying averaged spiking activity generated by the thousands of neurons surrounding the contact site of the electrode (Blenkmann et al., 2019; Lachaux et al., 2012; Ray & Maunsell, 2011; Rich & Wallis, 2017; Watson et al., 2018).

To eliminate any residual artefacts not rejected by visual inspection, responses with an amplitude larger than five standard deviations from the mean for more than 25 consecutive ms or with a power spectral density above five standard deviations from the mean for more than six consecutive Hz were excluded (this was applied to all activity types).

3.4 Surrogate Data Testing

Through robust statistical evaluation, surrogate data testing enables to ensure that the calculated measures are not obtained by chance (Lancaster et al., 2018). Since we employed information-theoretical principles, surrogate data testing is one of the most used methods in this project. Accordingly, we utilised it in most performed analyses. All resulting articles contain surrogate data approaches to assess the statistical significance of the respective information-theoretical variables. The realised surrogate tests tested for uncorrelated noise. That is to say, they statistically tested whether the respective information-theoretic quantity can be solely reduced to uncorrelated noise. This proceeded by repeatedly computing a measure of interest based on randomly shuffled neurophysiological data. These constructed values then build a null model distribution. To reject the null hypothesis that the measure of interest stems from uncorrelated noise, the true value should be far outside of any null model distribution (Fig. 2.4). To obtain a p-value, we then evaluated how many times the true value is greater or smaller than the surrogate values that build the null model distribution, divided by the total number of values (Ince et al., 2017; Lancaster et al., 2018; Timme & Lapish, 2018). This p-value was then directly used in Paper II and Paper III to compute significance ratios by simply dividing the number of significant variables by the total amount of considered variables (e.g., number of significant channels to the total amount of channels).

Chapter 4

Summary of Papers

In this chapter, the published work resulting from this project is summarised. Each summary gives the motivation behind the respective article and an overview of the underlying work. A short teaser of each paper is stated below, followed by a more detailed summary of each paper. All articles are displayed at the end of the thesis.

Paper I bundles much of the work and is about the implicit encoding of statistical structure within an auditory sequence. Specifically, the role of TPs as an implicit structure supporting the encoding of a random auditory stream is examined. This proceeds on the basis of information-theoretical principles and HFA originating from iEEG. In short, this work demonstrates how the brain continuously encodes structure in random stimuli involving a network outside of the auditory system, including hippocampal, frontal, and temporal regions.

Paper II elaborates on the application of the Encoded Information method proposed in Paper I. This article validated its use by applying it to both synthetic and real neurophysiological data and compared its efficiency to the MI measure.

Paper III is a continuation of Paper II, where the Encoded Information measure is further investigated. More precisely, based on iEEG recordings from humans and marmoset monkeys, the approach was applied to the activity types of theta, alpha, beta and evoked related potential (ERP) in addition to HFA. Further, its performance was compared to that of a conventional t-test, MI, Gaussian Copula MI and to the method of Neural Frequency Tagging.

4.1 Paper I: Direct Brain Recordings Reveal Continuous Encoding of Structure in Random Stimuli

Fuhrer, J., Glette, K., Ivanovic, J., Larsson, P. G., Bekinschtein, T., Kochen, S., Knight, R. T., Tørresen, J., Solbakk, A.-K., Endestad, T., & Blenkmann, A. (submitted). Direct brain recordings reveal continuous encoding of structure in random stimuli. *bioRxiv*. https://doi.org/10.1101/2021.10.01.462295

Motivation

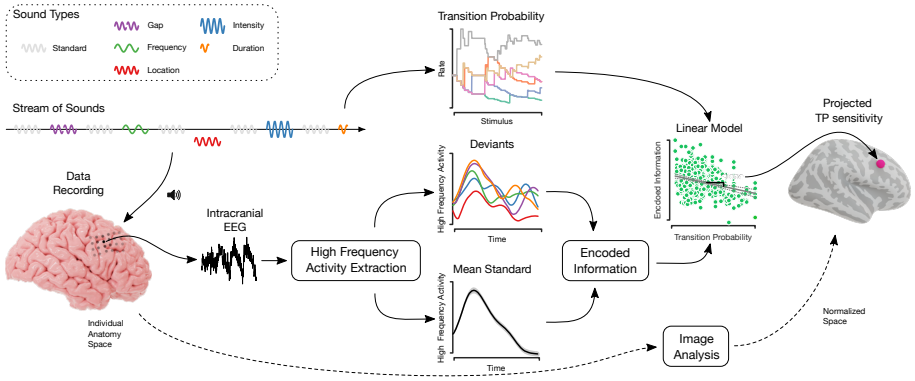
Detailed knowledge about where cortical mechanisms locate during implicit learning is limited since much of the work on statistical learning studies using neurophysiological methods deploy active tasks. Particularly, information about sub-cortical structures such as the hippocampus is limited. This was revealed when the core hypothesis of this thesis was conceptualised, stating that the brain is incessantly encoding patterns. Having access to iEEG with its low SNR, most of the project work invested in investigating these mechanisms through the introduced paradigm. Consequently, the present publication crystallises most of the conducted research. The desire was to reveal aspects of these mechanisms, accordingly contributing to the fundamental understanding of the brain.

Summary

At the beginning of this project part, well-established analysis of power or amplitude differences in the time series signals (usually evaluated with classic statistical hypothesis testing, e.g., t-tests) were first considered. Although there were some tendencies, it occurred that with these methods, no clear statement in support of the main hypothesis (cf., Chapter 1) could be made. Therefore other, possibly more sensitive, methods were sought. Approaches from information theory turned out to be the most promising, such that most analysis in this work is based on information-theoretical principles.

Individual trials, i.e., evoked responses to sounds, were then represented through algorithmic information theory, allowing to compare the responses that stem from standard responses to deviant responses. This part of the analysis took the role of validation for the suggested measure of encoded information, as there is quite a considerable amount of neuroimaging research on deviance detection or MMN. The superior temporal plane (comprising early auditory areas such as Heschl's gyrus) showed the highest sensitivity to deviant sounds relative to standards. This is followed by the posterior part of the insula, anterior cingulate cortices and the lateral part of the temporal cortex. Higher-order areas such as the pre-frontal cortex, on the other hand, showed less sensitivity to deviants. It was possible to confirm this trend by comparing the sensitivity distribution of deviants with a freely available anatomical hierarchy mapping. In short, the conducted analysis through algorithmic complexity is in line with the deviance detection literature.

(a) Overview of the analysis



(b) TP sensitivity across different brain areas

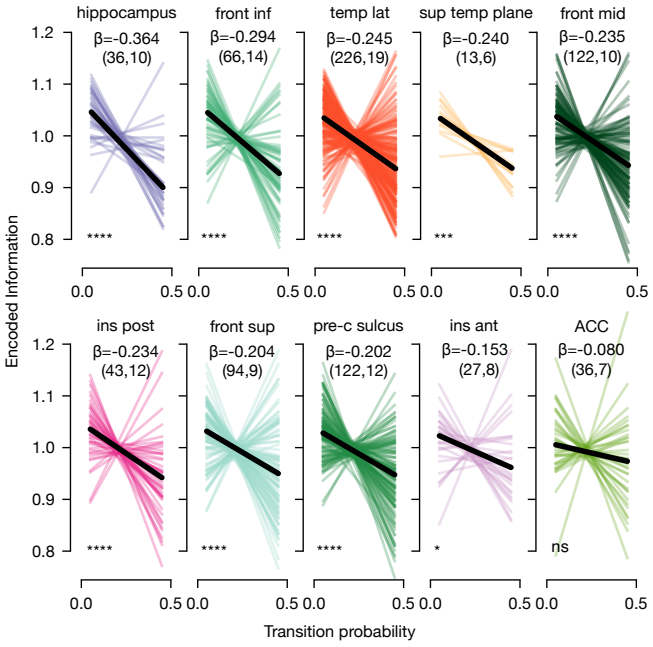


Figure 4.1: The two most essential figures from the article. **a**: A stream of sounds was presented to participants while recording their event-related electrical brain activity through intracranial electrodes. Emerging brain responses were treated through information-theoretical principles. Statistics were then extracted from the sequence and regressed with the information-based representation of brain responses. **b**: The main result of this work showing the sensitivity to TPs for individual regions during the implicit listening task.

The second part was then to investigate if any sensitivity to the perceptual variables can be found (cf. Chapter 2). Here, the main focus was on the TPs by performing a trial-by-trial analysis of all 22 subjects (Fig. 4.1a). Based on this analytic procedure, a picture was crystallised hinting at a sensitivity of the brain towards TPs in an implicit multi-dimensional listening task with random features. This emerged from a correlation analysis with HFA and TP estimates through robust linear regression models. It revealed that the subcortical structure of the hippocampus shows the highest sensitivity towards TP among all regions considered (Fig. 4.1b). This is an interesting finding, as its role is highly discussed in the scientific community, and would fit the idea that it is a rapid supramodal learner of arbitrary or higher-order associations in the sensory environment. The hippocampus is accompanied by the inferior frontal region. Its involvement during unattended tasks is less evident from prior reports but has been hypothesised to be required during complex computations (Conway, 2020). The regions are then followed by temporal regions, which suggests a network that involves both, higher-order areas and perceptual areas to encode the TPs. Overall, the findings of this article link the frameworks of statistical learning and predictive processing and are thus of great interest.

4.2 Paper II: Complexity-based Encoded Information Quantification in Neurophysiological Recordings

Fuhrer, J., Blenkmann, A., Endestad, T., Solbakk, A.-K., & Glette, K. (2022). Complexity-based encoded information quantification in neurophysiological recordings. *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2319–2323. https://doi.org/10.1109/EMBC48229.2022.9871501

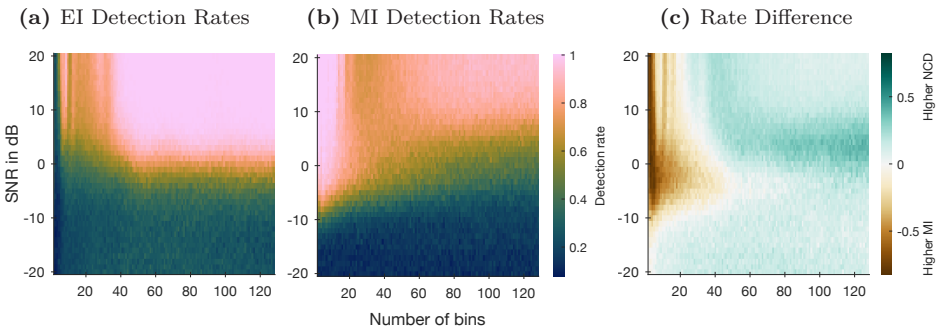


Figure 4.2: Main result from this paper. It shows detection rates over SNR for different numbers of bins for the measure of encoded information or MI. **a**: Mean detection rate across number of bins and SNRs for encoded information. **b**: The same results but for MI. **c**: Difference in the detection rates between encoded information and MI.

Motivation

One worthwhile analytic procedure that resulted from Paper I was the concept of Encoded Information. The former analysis considered a representation of a single trial (that is, single evoked responses). In this publication, the idea to draft a concept that allows us to compare two conditions based on algorithmic complexity, similar to that of a conventional t-test, followed. Consequently, the motivation was to compare mean responses on a channel basis. Therefore, the aim was to suggest a procedure for how to discriminate conditions based on algorithmic information theory.

Summary

To compute the encoded information variable of a channel, first, the idea of compressing the concatenation of all trials (all recorded evoked responses) for a specific condition was examined. For the iEEG dataset, per subject, there is 759.50 ± 360.85 number of trials for deviant responses and 715.14 ± 388.12 number of trials for standards. Consequently, the compression of such a long array turned out to be rather inefficient. For this reason, the procedure to compare the mean response stemming from the respective conditions was adopted. This procedure worked more robustly. The significance of the resulting encoded information variable was then assessed through surrogate data testing, by repeatedly shuffling trial responses across conditions (Fig. 2.4). Having established a general procedure, a parameter study and a comparison to the approach of MI were conducted. The main parameter of this method is the number of bins during the quantification procedure (section 2.5.2). To systematically examine this binning parameter, neurophysiological data was simulated through neural mass models, besides utilising the already treated iEEG recordings. The former enabled the generation of signals with a specific SNR such that a matrix of SNR to the number of bins combinations was generated, followed by an evaluation of both measures' performances by means of a detection rate. For both encoded information and MI, we defined the detection rate as the number of significant p-values to the total number of p-values resulting from the respective surrogate testing (assessing the statistical significance of encoded information or MI that discriminate experimental conditions). The outcome of this study was that a higher number of bins gives more robust detection rates for encoded information, whereas for MI, too high densities cause a decline in detection rate. In addition, both measures showed similar performance. However, the encoded information approach significantly detects more channels than MI. Especially for lower SNRs. This observation was confirmed by analysing the iEEG recordings. Accordingly, this information-theory-based procedure is an apt candidate to represent and investigate neurophysiological data.

4.3 Paper III: Quantifying Evoked Responses through Encoded Information

Fuhrer, J., Glette, K., Endestad, T., Solbakk, A.-K., & Blenkmann, A. (submitted). Quantifying evoked responses through encoded information. *bioRxiv*. https://doi.org/10.1101/2022.11.11.516096

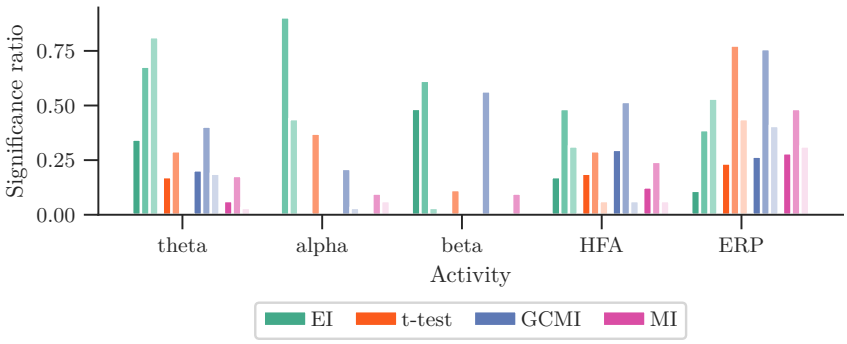


Figure 4.3: Significance ratios across activities for the different considered methods based on an iEEG roving oddball paradigm stemming from three marmoset monkeys.

Motivation

The relevance of this paper constitutes the broader investigation of the encoded information measure. So far, it only has been applied to HFA, allowing the question of whether it performs equally well on activity types other than HFA. Another point was to evaluate the methods on another data set and to compare the method to approaches other than MI. In short, the idea of this work was to possibly make a stronger claim about the value of encoded information. While this article solely focused on the technical aspect of comparing the deviant to the standard condition using different methods, the respective neuroscientific interpretation of the results was not part of this work.

Summary

The first data set we considered was the recording treated in Paper I and Paper II. Besides that, we investigated two additional data sets. The second data set was based on a roving oddball paradigm recorded from three marmoset monkeys through iEEG (Canales-Johnson et al., 2021; Komatsu et al., 2015). The third data set stemmed from iEEG electrodes located in the insular cortex, where participants (n=12, mean age 31.2 ±11.1 yr, 4 female) performed a verbal working memory task (vWM or recent-probes task; see Llorens et al., 2022).

First, we extracted theta (5 to 7 Hz), alpha (8 to 12 Hz), beta (12 to 24 Hz), and ERP activity from the recordings of the two data sets. In the case of the roving oddball paradigm, before that step, recordings were segmented and cleaned. Subsequently, besides the proposed approach, we applied the methods of conventional t-test, MI, advanced MI estimation through Gaussian Copula (GCMI; that is a different procedure as described in Chapter 3) and neural frequency tagging (NFT) to the data sets. NFT was included for the first data set because it constitutes a completely different approach compared to the others as it makes use of peaks in the power spectrum of the raw data that match a presentation rate of interest. However, because of the infrequent deviant presentations within the roving oddball paradigm, NFT could not be applied there. The approach of NFT results in one discrete value, indicating the sensitivity to deviant tones. The same applies to the encoded information. For the t-test and the GCMI, on the other hand, a course over all samples within the length of a trial (evoked response to a deviant or standard) is obtained. While it is then possible to make statements about time, the entire time series of the respective statistic also needs to be corrected for multiple comparisons across channels and trials. Based on 1078 channels from 22 subjects, encoded information revealed a greater number of sensitive channels than the t-test, MI and GCMI for alpha, beta and HFA activity. The t-test showed the highest significance ratio for the ERP activity, which is however not significant on a subject level. Interestingly, the NFT approach also identified a considerable channel sensitivity to deviant tone types. This was also confirmed by evaluating the methods across different brain areas. Further, results from the marmoset recordings suggest a similar performance of encoded information. In short, from all methods considered, encoded information and t-test showed the greatest number of channels with sensitivity to deviant sounds, whereas the former shows significantly better results for beta and HFA activity. Therefore, the two most effective methods were further investigated by examining how well the methods operate when only a limited amount of trials are available for channels with differing deviant sensitivity. Here, both measures showed similar effectiveness. Overall, the proposed procedure of encoded information proves as a worthwhile procedure. Among the methods considered, the encoded information can compete in detecting channels sensitive to deviating sound types across activities. Especially for beta and HFA activity, encoded information proves as a worthwhile method, detecting a higher number of sensitive channels in comparison to the other measures.

Chapter 5

Discussion

Besides outlining the conducted work in this thesis, I sketched the underlying background and methods in the former chapters. In the following, I discuss the work itself focusing on how the presented papers answer the research questions. Limitations and concepts for future research then follow.

5.1 How does the Brain Operate when Implicitly Confronted with an Unstructured Auditory Sequence?

Based on unstructured stimuli, manifested through random occurrences of tones deviating from a standard sound, this project was able to shed light on the basic underlying mechanisms that make use of statistical regularities. By demonstrating that the unattended brain shows sensitivity towards TPs, it was possible to infer that the brain inexhaustibly scans the environment for patterns presumably to continuously refine internal perceptual models. That draws on prior research suggesting that TPs constitute a "scaffold" on which higher-order processes operate (Dehaene et al., 2015). Besides illuminating this process, this project was also able to depict how individual brain regions are involved. The role of identified areas such as the hippocampus and inferior frontal cortices is highly debated in current research. Consequently, the insights of this thesis contribute to this discussion. While the hippocampal and inferior frontal cortices largely showed significant activity during attentive tasks, more and more research suggests that they could also play an essential role during unattended scenarios (Billig et al., 2022; Carbajal & Malmierca, 2018; Dürschmid et al., 2016; Frost et al., 2015; Schapiro et al., 2016; Williams, 2020).

5.1.1 The Brain Encodes TPs in Random Stimulus Sequences

On average, more frequent deviant transitions exhibited less encoded information in the recorded responses. Conversely, rarer transitions increased the information encoded in the brain responses. This finding demonstrates that events with low TP (i.e., less expected or more surprising events) possess higher information encoded in the deviant responses. Following the notion of predictive processing, the information encoded in each deviant response can be interpreted as a bottom-up prediction error signal. According to this, larger prediction errors emerge out of such an event, deriving from less accurate predictions. Consequently, higher cortical areas utilise the increased information to update the internal models. On the other hand, more frequent events (i.e., events with high TPs) are more expected and thus show less encoded information. Accordingly, the input to update the internal model is smaller as the prediction error decreases.

5.1.2 Involved Cortical Areas are Distributed

From all regions considered, the hippocampus contributes the most to temporal transition encoding between deviating events (Fig. 4.1b). In contrast to the other treated areas, the analysis results indicate that hippocampal responses exhibit high sensitivity to TPs while showing less sensitivity to deviant tones. This suggests that hippocampal activity may reflect a more generic context sensitivity to the events' probabilistic structure. In other words, it is more involved in learning about event occurrences within a given structure on an abstract level instead of encoding actual deviating events (Strange et al., 2005). Accordingly, new evidence for the role of the hippocampus during implicit learning is provided. This is in line with the recent hypothesis that the hippocampus is a rapid supramodal learner of arbitrary or higher-order associations in the sensory environment (Barascud et al., 2016; Billig et al., 2022; Conway, 2020; Covington et al., 2018; Daikoku, 2018; Frost et al., 2015; Henin et al., 2021; Schapiro et al., 2014, 2016; Thiessen, 2017; Turk-Browne et al., 2009; Williams, 2020).

Besides the hippocampus, another domain-general area showed sensitivity towards TP courses. Having a high sensitivity towards TPs and a low sensitivity to deviant tones, the inferior frontal cortex showed similar behaviour to that of the hippocampus. However, evidence of inferior frontal involvement in statistics-driven learning processes is sparse (Barascud et al., 2016; Conway, 2020; Williams, 2020). It mainly relies on explicit learning studies using fMRI (Huettel et al., 2002; Karuza et al., 2013). Consequently, this project provided supportive evidence for its role. This fits well with studies examining deviance detection in humans, where it is assumed to be a higher hierarchical node (Dürschmid et al., 2018; Phillips et al., 2015, 2016). Also, non-human primate iEEG studies manipulating the predictability of events showed a spatially dispersed contribution of regions that includes the prefrontal cortex in both passive auditory (Chao et al., 2018) and active visual paradigms (Bastos et al., 2020).

The third relevant region is a perceptual region. In contrast to the two preceding regions, the superior temporal plane showed a high TP sensitivity and a high deviant tone sensitivity. This suggests vital importance during the encoding of deviating auditory events and TPs between them. This is in line with previous reports stating a contribution to statistical learning (Daikoku, 2018; Karuza et al., 2013; McNealy et al., 2006; Meyniel & Dehaene, 2017; Williams, 2020). For this reason, perceptual processing of individual stimuli in lower hierarchical areas might be affected by the encoding of temporal patterns (Bell et al., 2016; Conway, 2020; Lu & Vicario, 2014; Yaron et al., 2012). Possible causes for this impact can be local processes, modulations from higher hierarchical areas or a combination of both.

In a recent iEEG study presenting 12 syllables within an auditory stream, Henin et al. (2021) observed that TPs are encoded in lower-order areas of the superior temporal plane and not in the hippocampus. The hippocampus showed sensitivity towards the unique representation of the identity (i.e., the specific higher-order chunk such as a word) within the treated sequences. Hence, hippocampal activity did not reflect involvement in forming the neural

representation of TPs but rather operations that built upon them. The main factor leading to this difference could be that the study from Paper I is based on passive listening with pure tones, while Henin et al. used active listening with syllables.

5.1.3 Main Contribution and Implications

Treating the research question at hand led to a comprehensive picture of neural correlates of statistical learning. Before, these correlates were bundled together from multiple studies (Conway, 2020; Daikoku, 2018; Williams, 2020). On an abstract level, the results point toward an architecture that requires both domain-general regions (such as the hippocampus or the inferior frontal cortex) as well as modality-specific brain regions (e.g., superior temporal plane). However, how exactly those regions interact remains to be investigated.

In this work, as far as is known, we give the first evidence from direct brain recording showing the encoding of statistical structures within an implicit listening task. One possible implication of the study is based on the implicitness or randomness within the paradigm. Assuming that the brain constantly infers statistics, studies need to be carefully designed. Especially since this mechanism also operates without any conscious intention. If it is not accounted for that, the brain might encode confounding patterns. This particularly comes to play if the control condition allows for such operations. Studies often depend on control conditions where random stimuli are used as if *no* relevant cortical processing occurs. As the present research suggests, that might be premature in some scenarios. Accordingly, differences between test and control conditions could be traced back to confounding pattern encoding, possibly engaging different brain areas. A theoretical example could be repeating cascades of descending or ascending sound sequences (similar to a musical scale) used to contrast a classic oddball paradigm (see Carbajal & Malmierca, 2018, Fig. 1b). Having a musical scale as a control possibly allows for the encoding of patterns. Either within one block, where one sound predicts the other (the TP matrix has ones on the diagonal above the main diagonal, else only zeros for the descending sequence) or across blocks (i.e., on a global scale, where one block constitutes one pattern). As a consequence, it should become the norm to consider the brain's sensitivity to statistical regularities (e.g., computing the TP matrix) for neuroimaging studies that employ sequences of repeating events. Especially for implicit learning paradigms, given that the attention-independent systems are presumably always active yet little neurophysiological evidence exists (Conway, 2020; Williams, 2020). Further, it might be possible that the encoding of statistics could serve as a plausible hypothesis to investigate. Especially for fields such as deviance detection or predictive processing where the consideration of perceptual variables is seldom, this variable could be worth examining.

5.1.4 Limitations

While Paper I thoroughly targets the first research question, it does not cover all implied aspects. The main finding of this thesis suggests an implicit learning process in which the brain internally infers TPs. As outlined above, this deduction is based on correlative but not causal argumentation. Hence, this conclusion needs to be backed up by prior research but fortunately fits well with the theoretical frameworks of statistical learning or predictive processing (Clark, 2013; Daikoku, 2018; Frost et al., 2019). Aside from that, for each time step, TPs were computed given all prior deviant tone occurrences. TPs courses, therefore, represent the theoretically optimal courses. However, it is more realistic that the brain's capability to integrate past tones is limited (Baumgarten et al., 2021; Maheu et al., 2019). Implementing such a mechanism for the analysis' TP estimation potentially allows us to obtain further insights. In this connection, the activation of the ROIs was also considered only over the entire stimuli. This step was deemed necessary to have as much statistical power as possible in the regression models. However, focusing only on specific phases along the time axis (for example, early or late phase) would be highly relevant to consider, especially for the two areas showing the highest sensitivity. It is plausible that the hippocampus, for instance, is only active during the initial phase, and once the internal model is established, it halts its contribution. That would fit well with the idea of a hippocampal role responsible for the learning of a more general or abstract context (Billig et al., 2022; Strange et al., 2005). Such an analysis would be even more insightful for the inferior frontal cortex. As outlined above, evidence of its involvement in statistics-driven learning processes mainly relies on implications from active exposure paradigms or on reports from sensory processing studies, where TPs are not considered. In this regard, connectivity measures that estimate the information flow between regions would be valuable. It would allow drawing a clearer picture of how different areas interact or how dynamic this network is. It is plausible that there is a phase (for instance the initial phase) with no sensitivity towards TPs. During this phase, statistics other than TP (such as IF) could be preferred (Maheu et al., 2019).

Besides that, the recordings originated from clinical patients suffering from drug-resistant epilepsy who were potential candidates for resective surgery. While epileptic activity was removed from the recordings, the data is not based on a healthy cohort. Additionally, the number and placement of electrodes were exclusively guided by clinical requirements. Consequently, the channel distribution is not uniform. While having already a considerable amount of subjects, one possibility to account for both aspects is to further increase the number of subjects.

5.1.5 Outlook

Following the present analysis, it is now possible to further explore underlying mechanisms and regional influences. Specifically, adding lower frequency bands to the analysis would enable disentangling their potential distinct roles in

information encoding and predictability signalling of sensory inputs. While alpha or beta activity could carry mechanisms of top-down predictions, it has rarely been directly identified. Focusing on the pre-onset sound interval of event response or examining cross-frequency coupling might help to reveal such predictive mechanisms. Also, considering temporal integration for TPs through a sliding window or information flow between regions is an interesting follow-up study. Comparing the TP sensitivity courses to anatomical hierarchy mappings could help to identify bottom-up or top-down driven processes as described in Bastos et al. (2015). That could potentially lead to identifying a dynamic hierarchy, namely functional hierarchy, where the hierarchical organization of brain areas dynamically changes given the demands of the task performed. Aside from that, statistics other than TPs could be added (such as IF, AF or simply time). Parts of this analysis have already been implemented but have not yet been evaluated. Among them is the analysis of lower frequency bands, multivariate regression models with multiple covariates or the sliding window approach of the TPs.

Further, non-adjacent TPs could be examined. That would be especially interesting in the case of an additional condition that involves attention. That leads to the next possible steps in terms of conducting future experiments. The natural choice would be to repeat the same experiment, only this time participants are instructed to pay attention to the sequence. That could enable an analysis of how the interplay between areas higher in the cortical hierarchy interacts with lower hierarchical brain regions. Results could then investigate how areas differ and whether this supports the hypothesis that two distinguished brain networks (implicit and explicit systems) exist (Conway, 2020). It is also possible to replace the pure tones with syllables or chunks of sounds that mimic natural speech to make the stimuli more realistic. That would possibly allow stronger links and increase the relevance to the study field of language acquisition (Conway, 2020; Dehaene et al., 2015; Henin et al., 2021; Vidal et al., 2019; Williams, 2020). Another worthwhile possibility is to (dynamically) vary the number of deviants within the sequence. If the brain would show sensitivity to this variation of deviants, direct learning of these pattern changes would be demonstrated. This could be further investigated by implementing individual blocks where, in each block, the TPs are manipulated differently. For example, one block with uniformly distributed TPs could be followed by a block with highly biased TPs. This could enable the investigation of how the active brain regions differ. If there is a dynamic network, it would be highly relevant to examine how this network changes throughout each block. Another option could be to use the same block several times, interspersed with randomly generated blocks. For the interspersed blocks there should be an increased activation in comparison to the learned regular block. However, it is also possible that it could be the other way around as recent research suggests (Barascud et al., 2016; Heilbron & Chait, 2018; Southwell & Chait, 2018). In short, based on these suggestions, it could be plausible to identify a general core network across the different scenarios.

5.2 To what extent is algorithmic information theory an apt approach to effectively represent neurophysiological signals?

Based on EEG, iEEG, MEG or fMRI, algorithmic information theory contributed to understanding aspects of consciousness (Canales-Johnson et al., 2020; Ruffini, 2017; Schartner et al., 2015, 2017; Sitt et al., 2014; Varley et al., 2020). From these studies, efficient approaches emerged to analyse brain states. Yet, these approaches are limited in their ability to directly quantify the level of similarity or disparity between brain states or conditions. With encoded information, this thesis addressed this issue by proposing an alternative measure grounded in algorithmic information theory able to compete with well-established methods. This novel approach was examined in detail in Paper II and compared to the popular measure of MI. Such a comparison was especially necessary because studies interested in quantifying the information contained in individual evoked brain responses often draw on the concept of Shannon entropy (i.e., classic information theory). This study concluded that encoded information exhibits an improved performance compared to MI. One possible explanation for this could be that encoded information is a compression-based approach. This mechanism allows compressing the entire time series while making use of structures along the time axis. That enables the exploitation of complementary structures within both input signals. Paper III revealed this approach to be particularly sensitive in detecting differences in beta band activity and HFA. For these signals, the measure of encoded information performs better, extending beyond the MI to the well-established analysis of power or amplitude differences in the time series (usually evaluated with classic statistical hypothesis testing, e.g., t-tests). For lower frequencies (theta activity) and ERP activity, encoded information's performance was similar to that of a power difference or MI.

In Paper II and Paper III we employed encoded information to discriminate experimental conditions from three different neurophysiological data sets. Two of these used auditory stimuli and one used visually presented stimuli. Central was the computation of a mean evoked response per condition given the recorded trials. Alternatively, it is also possible to directly compress the concatenation of trials (Fig. 2.4; Canales-Johnson et al., 2020; Sitt et al., 2014). However, this appeared to be problematic because of each trial's inherent noise. Thus, more trials lead to more noise contained within the concatenation. That causes a decrease in the compressor effectiveness merely based on the number of trials. By using the approach of computing a mean response from all trials, the noise contained across trials usually cancels out. That enables a more effective compression while the shuffling of trials between event types during significance testing still accounts for their variance. In this context, it is worth highlighting that current research debates the role of the so-called noise variability (Echeveste et al., 2020), which concerns the relevance of noise. That could be especially relevant for iEEG recordings with their high spectral resolution.

Paper I implemented the encoded information measure on a trial level. While the approach stays the same, the focus in this analysis was on how the information encoded in individual deviant trials differs concerning evoked responses to standard sounds. That is to say, principles based on algorithmic information theory reduced each deviant trial to one discrete value indicating how much it differs compared to standards. That vastly reduces the dimensions, which is especially helpful for high-dimensional data sets which exhibit a considerable amount of channels. Similar to MI, resulting values quantify the similarity of two signals. Accordingly, this representation is especially useful for subsequent regression analyses.

5.2.1 Limitations

While the property of making use of the entire time series is a key mechanism in the proposed measure, it is also one of its limitations. Because the compression procedure makes use of the temporal structure within the complete time course, it is difficult to obtain an encoded information time series. That would require compressing the values across trials for each time sample. Therefore, the advantage of exploiting the temporal structure would be omitted. Hence no information about the time course of responses is given. Besides that, signals need to be binned or simplified for the compressor to be efficient. Accordingly, some information is lost in that step. However, the chosen number of discrete steps during the binning procedure also has an effect here. Simulation results suggest that the performance of, for example, gzip stays about constant from approximately 32 to 128 bins. Another drawback is that, compared to the t-test, the statistical evaluation through surrogate data testing is more computationally demanding. However, choosing cluster-based permutation approaches instead of simple FDR or Bonferroni correction for the t-test demands similar computational power.

5.2.2 Outlook

Having applied this approach to iEEG recordings stemming from humans and other primates, the next natural choice is to evaluate the method's performance in terms of other neurophysiological recordings such as MEG or fMRI. Another option is to focus on the method by employing modern compressors, e.g., neural network compressors. Given the relations between algorithmic information theory and Shannon information theory, it also appears attractive to investigate a hybrid encoded information measure (e.g., using entropy estimation instead of compression). Another point is to offer more convenient accessibility of the source code by creating a toolbox compatible with popular toolboxes such as Fieldtrip (Oostenveld et al., 2011) or MNE (Gramfort et al., 2013).

In short, the proposed procedure grounded in algorithmic information theory is an apt candidate to represent and investigate neurophysiological data through information-theoretical principles. Encoded information offers a worthwhile alternative to conventional approaches. Especially for higher frequency activities,

encoded information showed a greater performance in comparison to well-established methods. Based on this, it can be concluded that algorithmic information theory is an apt candidate for the representation of neurophysiological signals. However, more research is needed to gain further insights into the performance of the encoded information measure.

Chapter 6

Conclusion

The main objective of this thesis was to investigate the hypothesis postulating that "*the brain continuously encodes structure within the sensory input stream*" (Chapter 1). Taken together, by making use of direct brain recordings, this project amply explored this hypothesis. In doing so, this thesis contributed to the understanding of brain functioning as well as to the methodology for analysing respective neurophysiological recordings.

By addressing the first research question, this thesis identified the brain's sensitivity towards the statistical regularity of TPs during an implicit listening task. That revealed the involvement of a network outside of the auditory cortex, including hippocampal, frontal and temporal regions. This neurophysiological evidence indicates an interactive network consisting of domain-general and modality-specific brain regions. Further, it supports the main hypothesis suggesting that the unattended brain incessantly encodes patterns attempting to make sense of the environment – even if they are random.

Treating the second research question, information-theoretical principles were leveraged. These methodological investigations explored the applicability of algorithmic information theory to analyse neurophysiological recordings. Investigations led to a worthwhile measure that efficiently discriminates conditions in neurophysiological recordings, affirming the present question.

While from this thesis viable insights emerged by considering the research questions, additional investigations are needed to make more precise statements. Besides a continuation of the analysis, the most important is the conduction of a follow-up experiment that involves attention and implements TPs differently. In addition, further examination of the encoded information measure by analysing different neurophysiological recordings is necessary.

Taken together, the findings of this thesis take a step forward in overcoming the present lack of neuroimaging studies investigating statistical structures such as TPs. TPs might constitute a central statistic for internal perceptual models at the core of predictive processing and statistical learning. In addition, this written thesis comprises only a part of the analysis conducted during this PhD project. The overall project established an interdisciplinary framework. Its particular composition of neuroscience and information theory yielded encouraging results. By further elaborating this framework in combination with novel experimental designs, it is worth continuing this work, facilitating the comprehension of the unknowns of the brain.

Bibliography

Arnal, L. H., & Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, *16*(7), 390–398. https://doi.org/10.1016/j.tics.2012.05.003

Atienza, M., L. Cantero, J., & Gómez, C. M. (1997). The mismatch negativity component reveals the sensory memory during rem sleep in humans. *Neuroscience Letters*, *237*(1), 21–24. https://doi.org/10.1016/S0304-3940(97)00798-2

Barascud, N., Pearce, M. T., Griffiths, T. D., Friston, K. J., & Chait, M. (2016). Brain responses in humans reveal ideal observer-like sensitivity to complex acoustic patterns. *Proceedings of the National Academy of Sciences*, *113*(5), E616–E625. https://doi.org/10.1073/pnas.1508523113

Bastos, A. M., Lundqvist, M., Waite, A. S., Kopell, N., & Miller, E. K. (2020). Layer and rhythm specificity for predictive routing. *Proceedings of the National Academy of Sciences*, *117*(49), 31459–31469. https://doi.org/10.1073/pnas.2014868117

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, *76*(4), 695–711. https://doi.org/10.1016/j.neuron.2012.10.038

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J.-M., Oostenveld, R., Dowdall, J. R., Weerd, P. D., Kennedy, H., & Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron*, *85*(2), 390–401. https://doi.org/10.1016/j.neuron.2014.12.018

Baumgarten, T. J., Maniscalco, B., Lee, J. L., Flounders, M. W., Abry, P., & He, B. J. (2021). Neural integration underlying naturalistic prediction flexibly adapts to varying sensory input rate. *Nature Communications*, *12*(1), 2643. https://doi.org/10.1038/s41467-021-22632-z

Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences*, *106*(5), 1672–1677. https://doi.org/10.1073/pnas.0809667106

Bell, A. H., Summerfield, C., Morin, E. L., Malecek, N. J., & Ungerleider, L. G. (2016). Encoding of stimulus probability in macaque inferior temporal cortex. *Current Biology*, *26*(17), 2280–2290. https://doi.org/10.1016/j.cub.2016.07.007

Billig, A. J., Lad, M., Sedley, W., & Griffiths, T. D. (2022). The hearing hippocampus. *Progress in Neurobiology*, 102326. https://doi.org/10.1016/j.pneurobio.2022.102326

Blenkmann, A. O., Collavini, S., Lubell, J., Llorens, A., Funderud, I., Ivanovic, J., Larsson, P. G., Meling, T. R., Bekinschtein, T., Kochen, S., Endestad, T., Knight, R. T., & Solbakk, A.-K. (2019). Auditory deviance detection in the human insula: An intracranial eeg study. *Cortex*, *121*, 189–200. https://doi.org/10.1016/j.cortex.2019.09.002

Canales-Johnson, A., Billig, A. J., Olivares, F., Gonzalez, A., Garcia, M. d. C., Silva, W., Vaucheret, E., Ciraolo, C., Mikulan, E., Ibanez, A., Huepe, D., Noreika, V., Chennu, S., & Bekinschtein, T. A. (2020). Dissociable neural information dynamics of perceptual integration and differentiation during bistable perception. *Cerebral Cortex*, *30*(8), 4563–4580. https://doi.org/10.1093/cercor/bhaa058

Canales-Johnson, A., Teixeira Borges, A. F., Komatsu, M., Fujii, N., Fahrenfort, J. J., Miller, K. J., & Noreika, V. (2021). Broadband dynamics rather than frequency-specific rhythms underlie prediction error in the primate auditory cortex. *Journal of Neuroscience*, *41*(45), 9374–9391. https://doi.org/10.1523/JNEUROSCI.0367-21.2021

Carbajal, G. V., & Malmierca, M. S. (2018). The neuronal basis of predictive coding along the auditory pathway: From the subcortical roots to cortical deviance detection. *Trends in Hearing*, *22*. https://doi.org/10.1177/2331216518784822

Chaitin, G. J. (1969). On the length of programs for computing finite binary sequences: Statistical considerations. *J. ACM*, *16*(1), 145–159. https://doi.org/10.1145/321495.321506

Chao, Z. C., Takaura, K., Wang, L., Fujii, N., & Dehaene, S. (2018). Large-scale cortical networks for hierarchical prediction and prediction error in the primate brain. *Neuron*, *100*(5), 1252–1266.e3. https://doi.org/10.1016/j.neuron.2018.10.004

Cilibrasi, R. L., & Vitányi, P. M. B. (2022). Fast phylogeny of sars-cov-2 by compression. *Entropy*, *24*(4). https://doi.org/10.3390/e24040439

Cilibrasi, R. L. (2007). *Statistical inference through data compression* [Doctoral dissertation, Institute for Logic, Language and Computation, Universiteit van Amsterdam]. Plantage Muidergracht 24, 1018 TV, Amsterdam, Holland.

Cilibrasi, R. L., & Vitányi, P. M. B. (2005). Clustering by compression. *IEEE Transactions on Information Theory*, *51*(4), 1523–1545. https://doi.org/10.1109/TIT.2005.844059

Clark, A. (2013). Whatever next? predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Conrad, K. (1958). *Die beginnende Schizophrenie: Versuch einer Gestaltsanalyse des Wahns*. Stuttgart: Thieme.

Conway, C. M. (2020). How does the brain learn environmental structure? ten core principles for understanding the neurocognitive mechanisms of statistical learning. *Neuroscience & Biobehavioral Reviews*, *112*, 279–299. https://doi.org/10.1016/j.neubiorev.2020.01.032

Covington, N. V., Brown-Schmidt, S., & Duff, M. C. (2018). The necessity of the hippocampus for statistical learning. *Journal of Cognitive Neuroscience*, *30*(5), 680–697. https://doi.org/10.1162/jocn_a_01228

Daikoku, T. (2018). Neurophysiological markers of statistical learning in music and language: Hierarchy, entropy, and uncertainty. *Brain Sciences*, *8*, 114. https://doi.org/10.3390/brainsci8060114

Daltrozzo, J., & Conway, C. M. (2014). Neurocognitive mechanisms of statistical-sequential learning: What do event-related potentials tell us? *Frontiers in Human Neuroscience*, *8*, 437. https://doi.org/10.3389/fnhum.2014.00437

Dehaene, S., Meyniel, F., Wacongne, C., Wang, L., & Pallier, C. (2015). The neural representation of sequences: From transition probabilities to algebraic patterns and linguistic trees. *Neuron*, *88*(1), 2–19. https://doi.org/10.1016/j.neuron.2015.09.019

Denham, S. L., & Winkler, I. (2020). Predictive coding in auditory perception: Challenges and unresolved questions. *European Journal of Neuroscience*, *51*(5), 1151–1160. https://doi.org/10.1111/ejn.13802

Domenech, P., & Dreher, J.-C. (2010). Decision threshold modulation in the human brain. *Journal of Neuroscience*, *30*(43), 14305–14317. https://doi.org/10.1523/JNEUROSCI.2371-10.2010

Dürschmid, S., Edwards, E., Reichert, C., Dewar, C., Hinrichs, H., Heinze, H.-J., Kirsch, H. E., Dalal, S. S., Deouell, L. Y., & Knight, R. T. (2016). Hierarchy of prediction errors for auditory events in human temporal and frontal cortex. *Proceedings of the National Academy of Sciences*, *113*(24), 6755–6760. https://doi.org/10.1073/pnas.1525030113

Dürschmid, S., Reichert, C., Hinrichs, H., Heinze, H.-J., Kirsch, H. E., Knight, R. T., & Deouell, L. Y. (2018). Direct evidence for prediction signals in frontal cortex independent of prediction error. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhy331

Echeveste, R., Aitchison, L., Hennequin, G., & Lengyel, M. (2020). Cortical-like dynamics in recurrent circuits optimized for sampling-based probabilistic inference. *Nature Neuroscience*, *23*, 1546–1726. https://doi.org/10.1038/s41593-020-0671-1

Falk, R., & Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review*, *104*(2), 301–318. https://doi.org/10.1037/0033-295X.104.2.301

Fischer, C., Morlet, D., Bouchet, P., Luaute, J., Jourdan, C., & Salord, F. (1999). Mismatch negativity and late auditory evoked potentials in comatose patients. *Clinical Neurophysiology*, *110*(9), 1601–1610. https://doi.org/10.1016/S1388-2457(99)00131-5

Foster, B. L., He, B. J., Honey, C. J., Jerbi, K., Maier, A., & Saalmann, Y. B. (2016). Spontaneous neural dynamics and multi-scale network organization. *Frontiers in Systems Neuroscience*, *10*. https://doi.org/10.3389/fnsys.2016.00007

Friston, K. (2008). Hierarchical models in the brain. *PLOS Computational Biology*, *4*, 1–24. https://doi.org/10.1371/journal.pcbi.1000211

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Friston, K. (2018). Does predictive coding have a future? *Nature Neuroscience*, *21*(8), 1019–1021. https://doi.org/10.1038/s41593-018-0200-7

Frost, R., Armstrong, B. C., & Christiansen, M. H. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin*, *145*(12), 1128–1153. https://doi.org/10.1037/bul0000210

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, *19*(3), 117–125. https://doi.org/10.1016/j.tics.2014.12.010

Fuhrer, J., Blenkmann, A., Endestad, T., Solbakk, A.-K., & Glette, K. (2022). Complexity-based encoded information quantification in neurophysiological recordings. *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2319–2323. https://doi.org/10.1109/EMBC48229.2022.9871501

Fuhrer, J., Glette, K., Endestad, T., Solbakk, A.-K., & Blenkmann, A. (submitted). Quantifying evoked responses through encoded information. *bioRxiv*. https://doi.org/10.1101/2022.11.11.516096

Fuhrer, J., Glette, K., Ivanovic, J., Larsson, P. G., Bekinschtein, T., Kochen, S., Knight, R. T., Tørresen, J., Solbakk, A.-K., Endestad, T., & Blenkmann, A. (submitted). Direct brain recordings reveal continuous encoding of structure in random stimuli. *bioRxiv*. https://doi.org/10.1101/2021.10.01.462295

Gilovich, T. (1991). *How we know what isn't so: The fallibility of human reason in everyday life.* Free Press.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, *17*(3), 295–314. https://doi.org/10.1016/0010-0285(85)90010-6

Graham, R. L., & Spencer, J. H. (1990). Ramsey theory. *Scientific American*, *263*(1), 112–117.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. S. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*(267), 1–13. https://doi.org/10.3389/fnins.2013.00267

Grünwald, P., & Vitányi, P. M. B. (2004). Shannon information and kolmogorov complexity. *CoRR*, *cs.IT/0410002*. https://doi.org/10.48550/arXiv.cs/0410002

Hahn, U., & Warren, P. A. (2009). Perceptions of randomness: Why three heads are better than four. *Psychological Review*, *116*(2), 454–461. https://doi.org/10.1037/a0015241

Heilbron, M., & Chait, M. (2018). Great expectations: Is there evidence for predictive coding in auditory cortex? [Sensory Sequence Processing in the Brain]. *Neuroscience*, *389*, 54–73. https://doi.org/10.1016/j.neuroscience.2017.07.061

Henin, S., Turk-Browne, N. B., Friedman, D., Liu, A., Dugan, P., Flinker, A., Doyle, W., Devinsky, O., & Melloni, L. (2021). Learning hierarchical sequence representations across human cortex and hippocampus. *Science Advances*, *7*(8). https://doi.org/10.1126/sciadv.abc4530

Higashi, H., Minami, T., & Nakauchi, S. (2017). Variation in event-related potentials by state transitions. *Frontiers in Human Neuroscience*, *11*, 75. https://doi.org/10.3389/fnhum.2017.00075

Huettel, S. A., Mack, P. B., & McCarthy, G. (2002). Perceiving patterns in random series: Dynamic processing of sequence in prefrontal cortex. *Nature Neuroscience*, *5*(5), 1546–1726. https://doi.org/10.1038/nn841

Ince, R. A., Giordano, B. L., Kayser, C., Rousselet, G. A., Gross, J., & Schyns, P. G. (2017). A statistical framework for neuroimaging data analysis based on mutual information estimated via a gaussian copula. *Human Brain Mapping*, *38*(3), 1541–1573. https://doi.org/10.1002/hbm.23471

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*(3), 430–454. https://doi.org/10.1016/0010-0285(72)90016-3

Karuza, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., & Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: An fMRI study. *Brain and Language*, *127*(1), 46–54. https://doi.org/10.1016/j.bandl.2012.11.007

Kikuchi, Y., Attaheri, A., Wilson, B., Rhone, A. E., Nourski, K. V., Gander, P. E., Kovach, C. K., Kawasaki, H., Griffiths, T. D., Howard, M. A., III, & Petkov, C. I. (2017). Sequence learning modulates neural responses and oscillatory coupling in human and monkey auditory cortex. *PLOS Biology*, *15*(4), 1–32. https://doi.org/10.1371/journal.pbio.2000219

Koelsch, S., Busch, T., Jentschke, S., & Rohrmeier, M. (2016). Under the hood of statistical learning: A statistical mmn reflects the magnitude of transitional probabilities in auditory sequences. *Scientific Reports*, *6*, 2045–2322. https://doi.org/10.1038/srep19741

Kolmogorov, A. N. (1968). Three approaches to the quantitative definition of information. *International Journal of Computer Mathematics*, *2*(1-4), 157–168. https://doi.org/10.1080/00207166808803030

Komatsu, M., Takaura, K., & Fujii, N. (2015). Mismatch negativity in common marmosets: Whole-cortical recordings with multi-channel electrocorticograms. *Scientific Reports*, *5*(1), 15006. https://doi.org/10.1038/srep15006

Lachaux, J.-P., Axmacher, N., Mormann, F., Halgren, E., & Crone, N. E. (2012). High-frequency neural activity and human cognition: Past, present and possible future of intracranial eeg research [High Frequency Oscillations in Cognition and Epilepsy]. *Progress in Neurobiology*, *98*(3), 279–301. https://doi.org/10.1016/j.pneurobio.2012.06.008

Lancaster, G., Iatsenko, D., Pidde, A., Ticcinelli, V., & Stefanovska, A. (2018). Surrogate data for hypothesis testing of physical systems. *Physics Reports*, *748*, 1–60. https://doi.org/10.1016/j.physrep.2018.06.001

Leonard, M. K., Bouchard, K. E., Tang, C., & Chang, E. F. (2015). Dynamic encoding of speech sequence probability in human temporal cortex. *Journal of Neuroscience*, *35*(18), 7203–7214. https://doi.org/10.1523/JNEUROSCI.4100-14.2015

Li, M., & Vitányi, P. (2008). *An introduction to kolmogorov complexity and its applications* (3rd ed.). Springer New York. https://doi.org/10.1007/978-0-387-49820-1

Li, M., Chen, X., Li, X., Ma, B., & Vitanyi, P. (2004). The similarity metric. *IEEE Transactions on Information Theory*, *50*(12), 3250–3264. https://doi.org/10.1109/TIT.2004.838101

Llorens, A., Bellier, L., Blenkmann, A., Ivanovic, J., Larsson, P., Lin, J., Endestad, T., Solbakk, A.-K., & Knight, R. (2022). Decision and response monitoring during working memory are sequentially represented in the human insula. *bioRxiv*. https://doi.org/10.1101/2022.10.25.513764

Lu, K., & Vicario, D. S. (2014). Statistical learning of recurring sound patterns encodes auditory objects in songbird forebrain. *Proceedings of the National Academy of Sciences*, *111*(40), 14553–14558. https://doi.org/10.1073/pnas.1412109111

Maguire, P., Moser, P., Maguire, R., & Keane, M. T. (2019). Seeing patterns in randomness: A computational model of surprise. *Topics in Cognitive Science*, *11*(1), 103–118. https://doi.org/10.1111/tops.12345

Maheu, M., Dehaene, S., & Meyniel, F. (2019). Brain signatures of a multiscale process of sequence learning in humans (F. de Lange & T. E. Behrens, Eds.). *eLife*, *8*, e41541. https://doi.org/10.7554/eLife.41541

Maheu, M., Meyniel, F., & Dehaene, S. (2022). Rational arbitration between statistics and rules in human sequence processing. *Nature Human Behaviour*, *6*(8), 1087–1103. https://doi.org/10.1038/s41562-021-01259-6

McNealy, K., Mazziotta, J. C., & Dapretto, M. (2006). Cracking the language code: Neural mechanisms underlying speech parsing. *Journal of Neuroscience*, *26*(29), 7629–7639. https://doi.org/10.1523/JNEUROSCI.5501-05.2006

Meyniel, F., & Dehaene, S. (2017). Brain networks for confidence weighting and hierarchical inference during probabilistic learning. *Proceedings of the National Academy of Sciences*, *114*(19), E3859–E3868. https://doi.org/10.1073/pnas.1615773114

Meyniel, F., Maheu, M., & Dehaene, S. (2016). Human inferences about sequences: A minimal transition probability model. *PLOS Computational Biology*, *12*(12), 1–26. https://doi.org/10.1371/journal.pcbi.1005260

Meyniel, F., Schlunegger, D., & Dehaene, S. (2015). The sense of confidence during probabilistic learning: A normative account. *PLOS Computational Biology*, *11*(6), 1–25. https://doi.org/10.1371/journal.pcbi.1004305

Mittag, M., Takegata, R., & Winkler, I. (2016). Transitional probabilities are prioritized over stimulus/pattern probabilities in auditory deviance detection: Memory basis for predictive sound processing. *Journal of*

Neuroscience, *36*(37), 9572–9579. https://doi.org/10.1523/JNEUROSCI.1041-16.2016

Näätänen, R., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity (mmn): Towards the optimal paradigm. *Clinical Neurophysiology*, *115*(1), 140–144. https://doi.org/10.1016/j.clinph.2003.04.001

Nasa. (2001). *Cydonia mensae*. National Aeronautics and Space Administration. Retrieved September 8, 2022, from https://mars.jpl.nasa.gov/gallery/atlas/cydonia-mensae.html

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). Fieldtrip: Open source software for advanced analysis of meg, eeg, and invasive electrophysiological data. *Computational intelligence and neuroscience*, *2011*, 156869. https://doi.org/10.1155/2011/156869

Paavilainen, P. (2013). The mismatch-negativity (mmn) component of the auditory event-related potential to violations of abstract regularities: A review. *International Journal of Psychophysiology*, *88*(2), 109–123. https://doi.org/10.1016/j.ijpsycho.2013.03.015

Parras, G. G., Nieto-Diego, J., Carbajal, G. V., Valdés-Baizabal, C., Escera, C., & Malmierca, M. S. (2017). Neurons along the auditory pathway exhibit a hierarchical organization of prediction error. *Nature Communications*, *8*(1), 2148. https://doi.org/10.1038/s41467-017-02038-6

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Learning in reverse: Eight-month-old infants track backward transitional probabilities. *Cognition*, *113*(2), 244–247. https://doi.org/10.1016/j.cognition.2009.07.011

Phillips, H. N., Blenkmann, A., Hughes, L. E., Bekinschtein, T. A., & Rowe, J. B. (2015). Hierarchical organization of frontotemporal networks for the prediction of stimuli across multiple dimensions. *Journal of Neuroscience*, *35*(25), 9255–9264. https://doi.org/10.1523/JNEUROSCI.5095-14.2015

Phillips, H. N., Blenkmann, A., Hughes, L. E., Kochen, S., Bekinschtein, T. A., Cam-CAN, & Rowe, J. B. (2016). Convergent evidence for hierarchical prediction networks from human electrocorticography and magnetoencephalography. *Cortex*, *82*, 192–205. https://doi.org/10.1016/j.cortex.2016.05.001

Piasini, E., & Panzeri, S. (2019). Information theory in neuroscience. *Entropy*, *21*(1). https://doi.org/10.3390/e21010062

Picton, T. W., John, M. S., Dimitrijevic, A., & Purcell, D. (2003). Human auditory steady-state responses: Respuestas auditivas de estado estable en humanos. *International Journal of Audiology*, *42*(4), 177–219. https://doi.org/10.3109/14992020309101316

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*, 79.

Ray, S., & Maunsell, J. H. R. (2011). Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. *PLOS Biology*, *9*(4), 1–15. https://doi.org/10.1371/journal.pbio.1000610

Rich, E. L., & Wallis, J. D. (2017). Spatiotemporal dynamics of information encoding revealed in orbitofrontal high-gamma. *Nature Communications*, *8*(1), 1139.

Rubin, J., Ulanovsky, N., Nelken, I., & Tishby, N. (2016). The representation of prediction error in auditory cortex. *PLOS Computational Biology*, *12*(8), 1–28. https://doi.org/10.1371/journal.pcbi.1005058

Ruffini, G. (2017). An algorithmic information theory of consciousness. *Neuroscience of Consciousness*, *2017*(1). https://doi.org/10.1093/nc/nix019

Saffran, J. R. (2020). Statistical language learning in infancy. *Child Development Perspectives*, *14*(1), 49–54. https://doi.org/10.1111/cdep.12355

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology*, *69*(1), 181–203. https://doi.org/10.1146/annurev-psych-122216-011805

Sarasa, G., Granados, A., & Rodriguez, F. B. (2019). Algorithmic clustering based on string compression to extract p300 structure in eeg signals. *Computer Methods and Programs in Biomedicine*, *176*, 225–235. https://doi.org/10.1016/j.cmpb.2019.03.009

Schapiro, A. C., Gregory, E., Landau, B., McCloskey, M., & Turk-Browne, N. B. (2014). The necessity of the medial temporal lobe for statistical learning. *Journal of Cognitive Neuroscience*, *26*(8), 1736–1747. https://doi.org/10.1162/jocn_a_00578

Schapiro, A. C., Turk-Browne, N. B., Norman, K. A., & Botvinick, M. M. (2016). Statistical learning of temporal community structure in the hippocampus. *Hippocampus*, *26*(1), 3–8. https://doi.org/10.1002/hipo.22523

Schartner, M., Pigorini, A., Gibbs, S. A., Arnulfo, G., Sarasso, S., Barnett, L., Nobili, L., Massimini, M., Seth, A. K., & Barrett, A. B. (2017). Global and local complexity of intracranial eeg decreases during nrem sleep. *Neuroscience of Consciousness*, *2017*(1). https://doi.org/10.1093/nc/niw022

Schartner, M., Seth, A., Noirhomme, Q., Boly, M., Bruno, M.-A., Laureys, S., & Barrett, A. (2015). Complexity of multi-dimensional spontaneous eeg decreases during propofol induced general anaesthesia. *PLOS ONE*, *10*(8), 1–21. https://doi.org/10.1371/journal.pone.0133532

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Sitt, J. D., King, J.-R., El Karoui, I., Rohaut, B., Faugeras, F., Gramfort, A., Cohen, L., Sigman, M., Dehaene, S., & Naccache, L. (2014). Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state. *Brain*, *137*(8), 2258–2270. https://doi.org/10.1093/brain/awu141

Solomonoff, R. (1964). A formal theory of inductive inference. part i. *Information and Control*, *7*(1), 1–22. https://doi.org/10.1016/S0019-9958(64)90223-2

Southwell, R., & Chait, M. (2018). Enhanced deviant responses in patterned relative to random sound sequences. *Cortex*, *109*, 92–103. https://doi.org/10.1016/j.cortex.2018.08.032

Strange, B. A., Duggins, A., Penny, W., Dolan, R. J., & Friston, K. J. (2005). Information theory, novelty and hippocampal responses: Unpredicted or unpredictable? *Neural Networks*, *18*(3), 225–230. https://doi.org/10.1016/j.neunet.2004.12.004

Thiessen, E. D. (2017). What's statistical about learning? insights from modelling statistical learning as a set of memory processes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160056. https://doi.org/10.1098/rstb.2016.0056

Thompson, S. P., & Newport, E. L. (2007). Statistical learning of syntax: The role of transitional probability. *Language Learning and Development*, *3*(1), 1–42. https://doi.org/10.1080/15475440709336999

Timme, N. M., & Lapish, C. (2018). A tutorial for information theory in neuroscience. *eNeuro*, *5*(3). https://doi.org/10.1523/ENEURO.0052-18.2018

Turk-Browne, N. B. (2012). Statistical learning in perception. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 3182–3185). Springer US. https://doi.org/10.1007/978-1-4419-1428-6_1707

Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience*, *21*(10), 1934–1945. https://doi.org/10.1162/jocn.2009.21131

Varley, T. F., Luppi, A. I., Pappas, I., Naci, L., Adapa, R., Owen, A. M., Menon, D. K., & Stamatakis, E. A. (2020). Consciousness & brain functional complexity in propofol anaesthesia. *Scientific Reports*, *10*(1), 2045–2322. https://doi.org/10.1038/s41598-020-57695-3

Vidal, Y., Brusini, P., Bonfieni, M., Mehler, J., & Bekinschtein, T. A. (2019). Neural signal to violations of abstract rules using speech-like stimuli. *eNeuro*, *6*(5). https://doi.org/10.1523/ENEURO.0128-19.2019

Vitányi, P. M. B., Balbach, F. J., Cilibrasi, R. L., & Li, M. (2009). Normalized information distance. In F. Emmert-Streib & M. Dehmer (Eds.), *Information theory and statistical learning* (pp. 45–82). Springer. https://doi.org/10.1007/978-0-387-84816-7_3

Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proceedings of the National Academy of Sciences*, *108*(51), 20754–20759. https://doi.org/10.1073/pnas.1117807108

Warren, P. A., Gostoli, U., Farmer, G. D., El-Deredy, W., & Hahn, U. (2018). A re-examination of "bias" in human randomness perception. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(5), 663–680. https://doi.org/10.1037/xhp0000462

Watson, B. O., Ding, M., & Buzsáki, G. (2018). Temporal coupling of field potentials and action potentials in the neocortex. *European Journal of Neuroscience*, *48*(7), 2482–2497. https://doi.org/10.1111/ejn.13807

Williams, J. N. (2020). The neuroscience of implicit learning. *Language Learning*, *70*(S2), 255–307. https://doi.org/10.1111/lang.12405

Yaron, A., Hershenhoren, I., & Nelken, I. (2012). Sensitivity to complex statistical regularities in rat auditory cortex. *Neuron*, *76*(3), 603–615. https://doi.org/10.1016/j.neuron.2012.08.025

Zhao, J., Hahn, U., & Osherson, D. (2014). Perception and identification of random events. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1358–1371. https://doi.org/10.1037/a0036816

Papers

Paper I

Direct Brain Recordings Reveal Continuous Encoding of Structure in Random Stimuli

I

Julian Fuhrer, Kyrre Glette, Jugoslav Ivanovic, Pål Gunnar Larsson, Tristan Bekinschtein, Silvia Kochen, Robert T. Knight, Jim Tørresen, Anne-Kristin Solbakk, Tor Endestad, Alejandro Blenkmann

Submitted, Biorxiv version: DOI: 10.1101/2021.10.01.462295.

Paper II

Complexity-based Encoded Information Quantification in Neurophysiological Recordings

Julian Fuhrer, Alejandro Blenkmann, Tor Endestad, Anne-Kristin Solbakk, Kyrre Glette

Appeared in *44rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2022*, DOI: 10.48550/arXiv.2205.01337.

II

Paper III

Quantifying Evoked Responses through Encoded Information

Julian Fuhrer, Kyrre Glette, Tor Endestad, Anne-Kristin Solbakk, Alejandro Blenkmann

In preparation, Biorxiv version: DOI: 10.1101/2022.11.11.516096.

III

Quantifying Evoked Responses through Encoded Information

Julian Fuhrer*[,a,b], Kyrre Glette[a,b], Anaïs Llorens[d], Tor Endestad[a,c,e], Anne-Kristin Solbakk[a,c,e,f], and Alejandro Blenkmann[a,c]

[a]*RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion, University of Oslo, Oslo, Norway*
[b]*Department of Informatics, University of Oslo, Oslo, Norway*
[c]*Department of Psychology, University of Oslo, Oslo, Norway*
[d]*Helen Wills Neuroscience Institute and Department of Psychology, University of California, Berkeley, USA*
[e]*Department of Neurosurgery, Oslo University Hospital, Rikshospitalet, Oslo, Norway*
[f]*Department of Neuropsychology, Helgeland Hospital, Mosjøen, Norway*

Abstract

Information theory is a viable candidate to advance our understanding of how the brain processes information generated in the internal or external environment. With its universal applicability, information theory enables the analysis of complex data sets, is free of requirements about the data structure, and can help infer the underlying brain mechanisms. The branch of algorithmic information theory seems particularly suitable since it can estimate the information contained in individual brain responses to different stimuli. Here, we propose a measure grounded in algorithmic information theory termed Encoded Information as a novel approach to analyze neurophysiological recordings. Specifically, it enables an assessment of the encoded information that brain responses share with one another by compressing the respective signals. By applying the approach to event-related potentials and event-related activity in different frequency bands originating from intracranial electroencephalography recordings of humans and marmoset monkeys, we demonstrate that the information-based encoding can compete well with conventional approaches such as the t-test or Mutual Information. Information-based encoding is attractive whenever one is interested in detecting where in the brain the neural responses differ across experimental conditions.

EEG | SEEG | ECOG | Information Content | Algorithmic Complexity | t-test | Frequency Tagging

Introduction

Efficient processing of *information* is a core capacity of the brain. It enables us to perceive rapidly, comprehend, identify changes, and engage with our environment – sometimes even without conscious effort. Accordingly, in attempting to understand the underlying brain mechanisms responsible for these capacities, it is useful to employ approaches that originate from *information* theory. This mathematical theory provides multivariate analysis tools, is not bound to a single type of data, is model-independent (i.e., does not require assumptions about the data itself) and can capture nonlinear interactions (1–4). Hence, those principles can be utilized for the analysis of neurophysiological recordings originating from scalp electroencephalography (EEG), intracranial EEG (iEEG), magnetoencephalography (MEG), or functional magnetic resonance imaging (fMRI).

Specifically, by measuring the degree of redundancy, *algorithmic* information theory (AIT) estimates the absolute information contained in individual brain responses. The higher the information content, the more complex its structure. Accordingly, the less compressible or more random the response. Assessing the absolute information content of responses recorded from different contact sites enables inferring the complexity of the activity recorded at the respective sensors and potentially identification of underlying dynamics.

While AIT has been applied to analyses of task-related cognitive operations or to discriminate between states of consciousness measured with EEG, iEEG, MEG, or fMRI recordings (5–9), its use in

*E-mail address: julianpf@uio.no

neuroscience has been limited despite its clear potential. Here, we demonstrate the measure of *encoded information* (EI; 10) as an advantageous tool to directly quantify the level of similarity between responses of neurophysiological data across experimental conditions. By use of iEEG recordings stemming from 34 humans (11, 12) and three marmoset monkeys (13, 14) which were exposed to passive and active paradigms as well as auditory or visual stimuli, we validate this approach of encoded information considering theta (5 to 7 Hz), alpha (8 to 12 Hz), beta (12 to 24 Hz), high-frequency (HFA; 75 to 145 Hz) and broadband event-related potential (ERP) activity by comparing it to that of a conventional t-test, to Mutual Information (MI), Gaussian Copula Mutual Information (GCMI; 3), and to Neural Frequency Tagging (NFT; 15, 16).

Material & Methods

Test Paradigms and Neurophysiological Recordings

To evaluate our proposed method, we examined its sensitivity to discriminate experimental conditions from three different neurophysiological data sets. Two of these used auditory stimuli and one used visually presented stimuli. We focused our analysis on the cortical activity types of theta, alpha, beta, HFA, and ERP activity.

Extraction of Brain Activities

To obtain the ERPs, data were low-pass filtered at 30 Hz using a sixth-order Butterworth filter. Theta, alpha, and beta frequency bands were extracted from the demeaned signals using wavelet time-frequency transformation (Morlet wavelets) based on convolution in the time domain (17). All trials were then baseline corrected by subtracting the mean amplitude of the baseline period of each trial and frequency band from the entire trial (see respective sections for the different data sets for the used baseline intervals).

To extract the HFA, the pre-processed data were filtered into eight bands of 10 Hz ranging from 75 to 145 Hz by use of bandpass filters. Next, the instantaneous amplitude signal of each filtered signal was computed by applying a Hilbert transform to the filtered time series leading to the analytic signal (18), constituting a complex-valued time series. The analytic amplitude time series or signal envelope corresponding to specific frequency bands was then obtained by applying Pythagoras' Theorem. To obtain one time series across all eight frequency bands, their mean amplitude value was calculated. As the last step, the respective time series were normalized by dividing them by a mean baseline period which was computed from all trial recordings. This resulted in a normalized measure that was relative to the baseline activity and termed HFA.

To eliminate any residual artifacts not rejected by visual inspection, responses with an amplitude larger than five standard deviations from the mean for more than 25 consecutive ms, or with a power spectral density above five standard deviations from the mean for more than six consecutive Hz were excluded.

Optimum-1 Paradigm

We analyzed experimental iEEG data obtained from intracranial electrodes implanted in (self-reported) normal-hearing adults with drug-resistant epilepsy. Analyses of this data have been previously presented in (5, 10, 11). Participants (n=22, mean age 31 years, range 19 to 50 years, 6 female) performed a passive auditory oddball paradigm where a standard tone alternated with random deviant tones. The tones had a duration of 75 ms and were presented every 500 ms in blocks of 5 min consisting of 300 standards and 300 deviants. At the beginning of each block, 15 standards were played. To capture automatic, stimulus-driven processes, participants were asked not to pay attention to the sounds while reading a book or magazine. They completed 3 to 10 blocks, providing at least 1800 trials (for details, see 10, 11). From the 22 participants, a total of 1078 channels (mean: 48, range: 12 to 104) were available after data cleaning. Data were then segmented into 2000 ms epochs (750 ms before and 1250 ms after tone onset) and demeaned. The different activity types were then extracted, and for each activity type, differences between standard and deviant tone responses were evaluated in the 400 ms time window following the

sound onset across channels and subjects. The baseline window was from −100 to 0 ms relative to tone onset. Additionally, for this data set the neural oscillation synchronization to the tone onset frequency (2 Hz) and the frequency for standard and deviants tones (1 Hz) were examined by considering the pre-processed data (needed for the NFT approach).

Roving Oddball Paradigm

Further, we considered experimental iEEG data collected during a passive auditory roving oddball paradigm on three awake adult male common marmosets (*callithrix jacchus*). This experimental data has been previously studied in (see 13, 14). The paradigm consisted of trains of three, five, or 11 repeated single tones of 20 different frequencies (250 to 6727 Hz with intervals of ¼ octaves). All tones were pure sinusoidal tones, lasted 64 ms (7 ms rise/fall) and there was stimulus onset asynchrony 503 ms between them. For each sound train, all tones were identical but varied across tone trains. Consequently, the mismatch occurred between the transition from one train to another in the form of a frequency change. Accordingly, the last tone of a train is defined as a standard tone, while the first tone of a new train is considered a deviant tone. Standard to deviant tone transitions then occurred 240 times during a recording session.

The number of implanted electrodes varied from monkey to monkey. For monkey "Fr", 32 channels were implanted in the left hemisphere epidural space, for "Go", 64 channels were implanted in the right hemisphere, and for monkey "Kr", 64 electrodes were implanted in the right hemisphere (Fig. 6c; see 13, 14, for more detailed information). Recorded data were re-referenced through an average reference montage and epoched into −950 to 2000 ms segments relative to the standard or deviant tone onset. The different activity types were then extracted and baseline corrected (by use of the −100 to 0 ms time interval relative to tone onsets). For the analysis, all recordings were shortened to the −100 to 350 ms interval relative to sound onsets.

Verbal Working Memory Paradigm

As a third data set, we investigated iEEG data stemming from the insular cortex during a verbal working memory task (vWM; see 12). Participants (n=12, mean age 31.2 ±11.1 yr, 4 female) performed a recent-probes task, where in each trial a list of five letters was displayed (stimulus duration 500 ms, inter-onset interval 1000 ms) on the computer screen. The letter presentation was followed by 4 s maintenance and a 2 s probe period. During the latter, a probe letter was displayed where the participants had to answer whether the presented probe letter was in the current list (p=0.5). In total, 144 trials were presented to each participant in a pseudo-random order within three blocks (each 10 min).

From the twelve participants, a total of 90 bipolar channels (mean: 7.5, range: ±5.9) were available after data cleaning. Data were then segmented into 16 s epochs (−12 s before and 4 s after probe period) and demeaned. The different activity types were then extracted, and for each activity type, differences between maintenance (−2 to 0 s) and probe period (0 to 2 s) were evaluated across channels and participants. The window for the baseline correction was from −9.5 to −8.5 s, i.e., the second preceding the presentation of the letter list.

Encoded Information

We estimated the EI of the mean responses of the different experimental conditions as described in (10). In short, by employing algorithmic or Kolmogorov Complexity (K-complexity), we estimated the EI between conditions through the measure of Normalized Compression Distance (NCD) (2, 19). For a pair of signals (x, y), it is defined as

$$\mathrm{NCD}(x, y) = \frac{C(xy) - \min(C(x), C(y))}{\max(C(x), C(y))},$$

with $C(xy)$ denoting the compressed size of the concatenation of x and y, and $C(x)$ and $C(y)$ their respective size after compression (2, 19). Further, the NCD is non-negative, that is, it is $0 \leq \mathrm{NCD}(x, y) \leq 1 + \epsilon$, where the ϵ accounts for the imperfection of the employed compression technique. Small NCD values suggest similar signals, and high values indicate rather different signals.

To obtain a compressed version of a signal, it was first simplified by grouping its values into 128 discrete steps (bins). The bins covered equal distances and in a range between the global extrema of all the signals considered. The compressor received the indices of the bins that contained the elements of the signal (6, 7). Compression then proceeded through a compression routine based on Python's standard library with gzip. The statistical hypothesis testing was then performed through a permutation-based approach as described below.

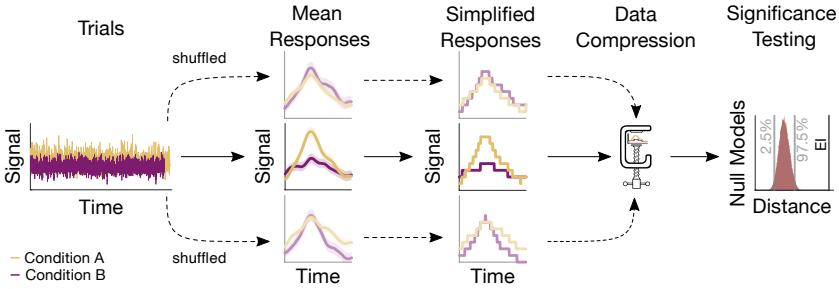


Figure 1: Sketch of the procedure for an example electrode. Based on the trials, a mean response for each condition is computed. Trials are then shuffled, resulting in surrogate mean responses. Subsequently, these signals undergo a simplification procedure, followed by their compression. The output of the compression routine is the EI, quantifying the similarity between responses. The resulting values are then evaluated, leading to a null model distribution. This distribution serves to assess the significance of the actual EI value.

t-test

The t-test is one of the most common methods to compare amplitude or power time series mean value differences across experimental conditions. For each sample and channel, a two-sided t-test was performed, where the resulting t-value represented the activity difference between the two conditions (e.g., standard or deviant). To correct for multiple comparisons across samples and channels, a False Discovery Rate (FDR) adjustment was applied with an FDR of 0.05.

Neural Frequency Tagging

For the optimum-1 data, we used NFT to identify the brain's capability to automatically segment the continuous auditory stream (15, 16). In the respective auditory sequence, the two main segments are represented by the frequency of sound onsets (2 Hz) or by the frequency of transitions between standards to standards or deviants to deviants (i.e., half the frequency or 1 Hz). If the neural activity of a recording site showed such "tagging" of frequency-specific properties within the stimuli, we defined it as "responsive". If it tagged half the frequency, we identified it as sensitive to the pattern of a standard-deviant alternation.

To assess this tagging, we computed the power-spectral density (PSD) of the epoched raw data using Welch's method (Fig. 2; 20). Subsequently, we estimated the signal-to-noise ratio (SNR) of the PSD (21). Here, SNR defined the ratio of power in a given frequency (signal) to the average power in the surrounding frequencies (noise). By doing so, we normalized the spectrum and accounted for the $1/f$ power decay (20). We then identified significant peaks through a lower threshold consisting of two times the standard deviation. The latter was estimated by computing the median absolute deviation, which was obtained by taking the median SNR multiplied by the constant distribution-dependent scale factor (Fig. 2). In the case of normally distributed observations, it reflects the 50 % of the standard normal cumulative distribution function, leading to a scale factor of 1.4826 (22, 23).

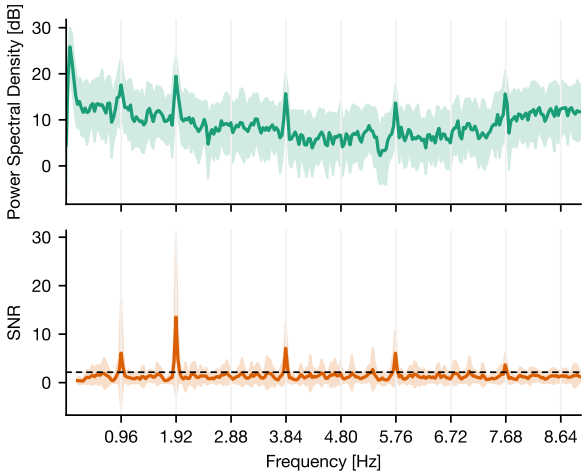


Figure 2: PSD and SNR of a responsive channel located in the superior temporal sulcus of a human. The activity of the channel shows synchrony to the main frequency of tone onsets, which is around 2 Hz. The displayed deviation is due to a constant lag of tone presentation during recording (resulting theoretical main sequence is at 1.919 Hz). Further, it shows synchrony to half the main frequency indicating that the underlying region of this channel tags the presentation rate of solely standard or deviant tones. Thus, it discriminates between standard and deviant conditions. The dashed line indicates the statistical threshold.

Mutual Information

Besides the t-test and NFT, EI was compared to the measure of MI. While it also grounds in information theory, in contrast to EI it draws on the concept of Shannon entropy (i.e., classic information theory). For a discrete random variable x with N outcomes, the entropy can be defined as

$$H(x) = -\sum_{k=1}^{\mathrm{N}} \mathrm{p}(x_\mathrm{k}) \log \mathrm{p}(x_\mathrm{k}) \,, \tag{1}$$

with $\mathrm{p}(x_\mathrm{k})$ being the occurrence probability for each element $x_\mathrm{k}, \ldots, x_\mathrm{N}$ of x. Given this definition, the MI between two discrete random variables (x, y) with N or M outcomes can be defined as

$$\begin{aligned} \mathrm{MI}(x; y) &= H(x) - H(x|y) \\ &= H(y) - H(y|x) \\ &= \sum_{\mathrm{k}=1}^{\mathrm{N}} \sum_{\mathrm{j}=1}^{\mathrm{M}} p(x_\mathrm{k}, y_\mathrm{j}) \log \frac{p(x_\mathrm{k}, y_\mathrm{j})}{p(x_\mathrm{k}) p(y_\mathrm{j})} \,, \end{aligned} \tag{2}$$

with the joint probability $p(x_\mathrm{k}, y_\mathrm{j})$ and the marginal probabilities $p(x_\mathrm{k})$ and $p(y_\mathrm{j})$.

This estimation of MI requires estimating the probability distribution of Eq. 2 by binning the signal into discrete steps. This is followed by a maximum likelihood estimation, yielding the probability distribution estimates. In our analysis, the respective signals were binned into four bins (1, 3).

Besides this binning approach, a novel estimation technique of MI after Ince et al. (3) was employed. With this approach, MI is estimated via Gaussian Copula. In short, each univariate marginal distribution is transformed into a standard normal. Subsequently, a Gaussian parametric MI estimate is applied. That yields a lower bound estimate of MI, named GCMI. Statistical significance testing was then estimated through surrogate data testing.

Surrogate Testing

The statistical significance of the information-based measures (EI, MI, and GCMI) was assessed through surrogate data testing. Accordingly, p-values were obtained by evaluating the observed information-based quantity in terms of a null distribution (Fig. 1). For EI and binned MI, null distributions were created by repeatedly shuffling the trials (i.e., single evoked responses) between conditions (e.g. standard and deviant) and then re-computing the information-based measure. For the GCMI, this proceeded for each sample of the time series signal. Single-sided p-values lower or equal to 0.05 were considered statistically significant. To correct for multiple comparisons across channels, FDR adjustment was applied with an FDR of 0.05.

Significance Ratio

To compare the different methods, the ratio of significant channels to the total amount of channels was determined. This was computed either for each subject, where the total significance ratio was the mean ratio across all subjects accompanied by a bootstrapped 95 % confidence interval. Or by taking the total ratio by collapsing across all channels (regardless of subjects). The latter was used when considering individual brain regions, where the number of channels was limited, leading to distorted ratios with large confidence intervals. For the roving oddball data, the ratio of each monkey was considered individually.

Results

Based on three neurophysiological data sets, we compared our method's performance on discriminating evoked responses to that of a conventional t-test, MI, and GCMI across five different cortical activity types. Additionally, we compared it to NFT. Overall, EI and t-test showed the greatest significance ratios (Fig. 3a, 6a, and 7a).

Across the 1078 channels stemming from the optimum-1 paradigm, for alpha, beta, and HFA activity, EI's significance ratio was significantly greater than for all methods considered (Fig. 3a) while the t-test showed the greatest significance ratio for the ERP activity (EI and t-test comparison on a subject level: two-sided paired t-tests, p=[1.43e−3, 3.61e−2, 2.87e−3 and 6.44e−3]). MI resulted in the lowest ratios, although it was only significantly lower on a subject level for HFA (two-sided paired t-test, p_{min}=3.02e−2). Note that for human iEEG, the number of channels and their distribution is not constant across subjects and can thus vary greatly. Consequently, the fewer channels for a subject, the less robust the estimated individual significance ratio.

The observations from this data set are in line with the results from the marmoset recordings, which had 128 channels in total. Based on the significance ratio of each marmoset monkey, for theta, alpha, and beta activity, EI detected more channels than the other approaches. For HFA, GCMI and EI performed similarly well, while for ERP, GCMI and t-test performed best (Fig. 6a). Additionally, common to all methods was the variability in significance ratios across monkeys.

A similar observation was made by considering EI's performance for the vWM task (Fig. 7a). Across all activities, EI showed the highest significance ratios. Interestingly, for the ERPs, both EI and GCMI had a significantly greater significance ratio than the t-test (p_{min} = 3.80e−2). Besides that, especially GCMI exhibited a consistently higher significance ratio across activities in comparison to the optimum-1 and roving oddball data sets.

To assess channels that the methods commonly detected, we divided the absolute numbers of significant channels by the total amount of significant channels for each pairwise method combination (Fig. 6c, i.e., each method's specific unique set of significant channels plus their intersection) and defined this ratio as the intersection ratio. For the optimum-1 paradigm, the intersection ratio between the t-test and GCMI was the greatest across all activity types, reaching its maximum for the ERP activity (Fig. 3b). For beta and HFA activity, there was also a high intersection ratio between the t-test and GCMI. However, for these activities, the number of significant channels was low (Fig. 3a), which led to higher percentages. This can also be observed by correlating the method-specific channel distributions across subjects (Fig. 3c), i.e., by correlating the significance ratios across subjects for each method-to-method combination. Besides

that, the approaches correlated the least for beta activity, which was also the activity with the lowest significance ratios. Overall, there was a high correlation across methods (Fig. 3c).

The relatively high intersection ratio between the t-test and GCMI across activities was not observable for the roving oddball paradigm and the vWM task. For the former, the intersection ratios were the highest for HFA and ERP activity (Fig. 6b), while for the latter it was highest for ERP (Fig. 7b&7c). Furthermore, the same effect as for the optimum-1 paradigm recordings occurred for the roving oddball paradigm. For HFA, the monkey Fr showed maximum intersection ratios between t-test, GCMI, or MI. Of note, the respective significance ratios were rather low (Fig. 6a). Moreover, when comparing Kr to Go it appeared that the latter had a smaller number of significant and overlapping channels for EI and t-test. However, because Kr exhibited more significant channels for both methods, the intersection ratio was of comparable magnitude for both monkeys.

For the optimum-1 paradigm data set, we further evaluated the NFT of all contact sites. The number of responsive channels, i.e., the number of channels that solely tagged the frequency of standards or deviants, was lower in comparison to all other methods (two-sided paired t-test between NFT and MI, $p=6.87e{-}3$).

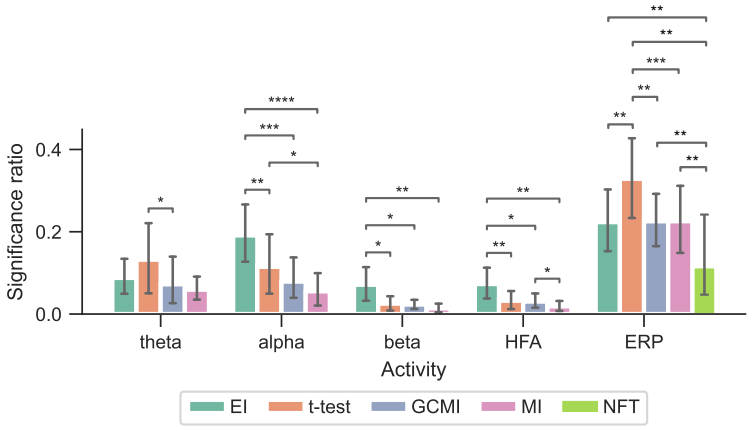
Given the differing sensitivity to deviating sounds across individual brain regions and the number of participants for the optimum-1 paradigm, we then compared the methods' significance ratios across individual cortical areas. For this reason, we evaluated the methods' performances across different areas comprising temporal, frontal, insular, peri-central sulci, and anterior cingulate cortices (ACC), as well as the hippocampus (Fig. 4; see 5, for exact definitions). Overall, the significance ratios of EI aligned with the other methods across brain regions and activity types. Specifically, the ratio for higher cortical areas such as the superior or middle frontal cortex was low while being higher in areas such as the superior temporal plane (which includes Heschl's gyrus). The NFT method together with MI showed the lowest sensitivities across brain regions. However, there were higher or equally high ratios in responsive brain areas such as the superior temporal plane or posterior insular cortices.

To examine the capabilities of the two best-performing methods, we further applied the approach of EI and t-test to scenarios where only a reduced number of trials were available. The number of trials varied from 1 to 100 % of all available trials (759.50 ± 360.85 trials for deviant responses and 715.14 ± 388.12 trials for standards across all channels). For each percentage and condition, trials were randomly chosen. This step was repeated 50 times for each trial increment to obtain an average p-value as a function of the number of trials (Fig. 5). Channels were selected from the optimum-1 paradigm with differing sensitivity to deviating tones. The channels were located in the respective 25, 50, 75, and 97.5-percentiles of the t-value distribution emerging from HFA. As can be seen in Fig. 5, the two methods only marginally differed in reaching statistical significance, although EI tended to reach this threshold slightly earlier.

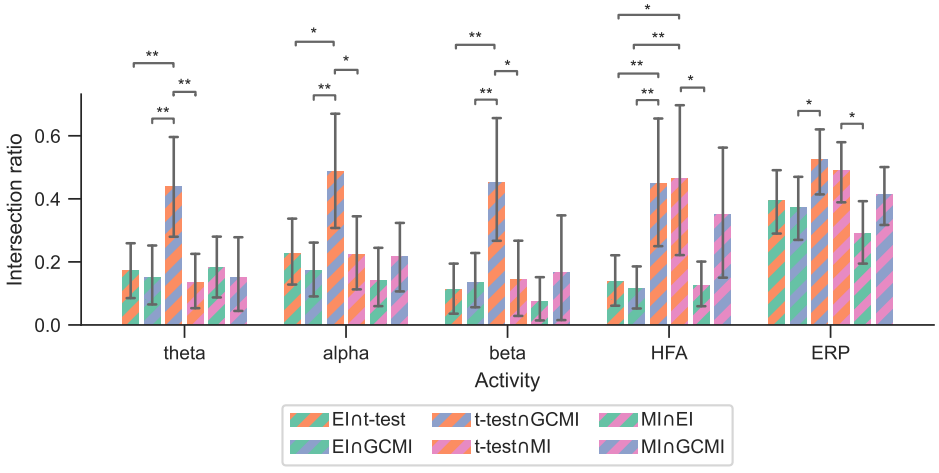
Discussion

Despite its ability to estimate the absolute information contained in individual brain responses, the use of AIT in neuroscience is limited. Hence, information-theoretical approaches in neuroscience are needed to facilitate its use. With the measure of EI, we proposed information-based task condition discrimination as an alternative to the classical t-test approach. Using compression as the core principle, this procedure quantifies the similarity between recordings stemming from different conditions. By applying this procedure to five cortical activity types and comparing it to that of t-test, MI, GCMI, and NFT, we demonstrate the high sensitivity of EI.

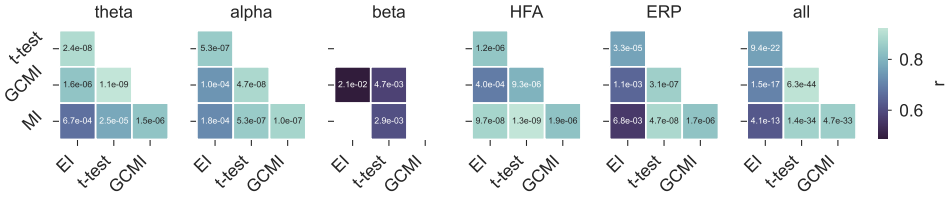
Besides their individual procedures in estimating test statistics, the examined methods also differ in how the data were processed. While for the t-test and GCMI approach, each sample or time point across trials is considered independently, EI, and also MI in its current implementation make use of all time samples by computing the respective mean responses across trials. Consequently, for each channel, the first two methods result in test statistics along the time axis, while the last two output one concrete channel-dependent test statistic. Such a time course is advantageous when the core interest is within the time domain or latency of responses, i.e., at which time point responses differ most. However, when it comes to assessing how sensitive contact sites are towards different conditions, this time course is of a minor



(a) Significance ratios across activities and methods.



(b) Intersection of significant channels across methods.



(c) Correlation of methods' performances across subjects.

Figure 3: Performance of the different methods. The error bars indicate the 95% CIs across subjects. Importantly, each subject has a unique electrode distribution such that the range of significant channels can greatly vary. **a**: Significance ratio across electrophysiological activity types. Statistical significance is indicated with * $p \leq 5e{-}2$, ** $p \leq 1e{-}2$, *** $p \leq 1e{-}3$ and **** $p \leq 1e{-}4$. **b**: Intersection of the significant iEEG channels for each method combination. Each number is shown relative to the total number of significant channels. **c**: Correlation matrices comparing the subject-specific significant ratios. The respective p-value is annotated in each square.

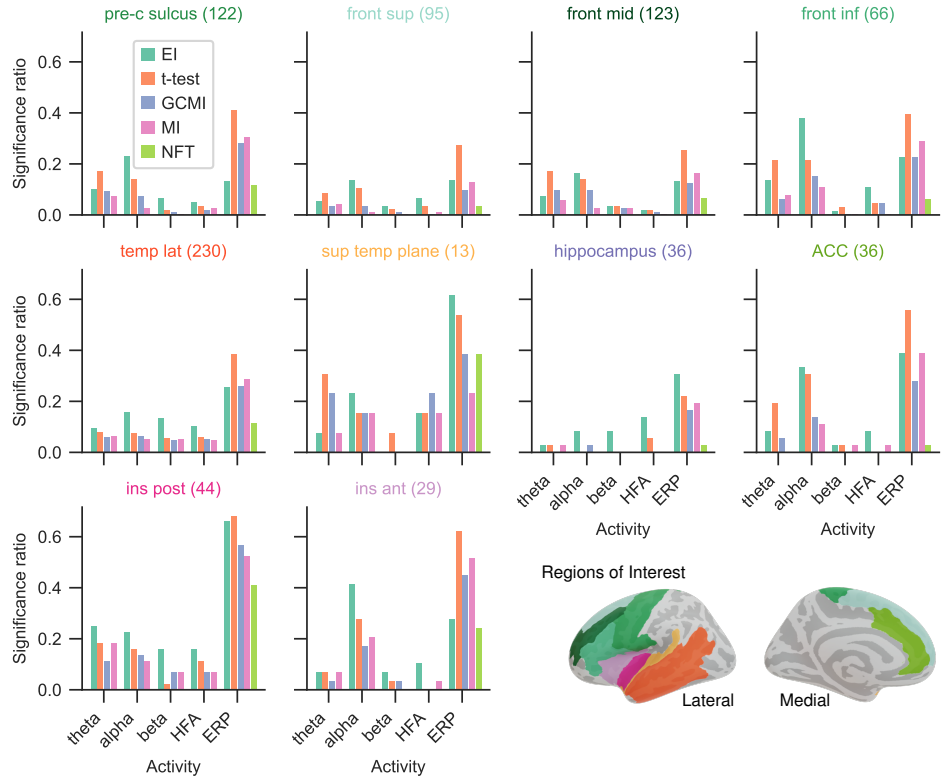


Figure 4: Significance ratio across brain regions for the different methods. The number of channels per ROI is indicated in brackets (For more detailed information on the ROIs, see supplementary of 5).

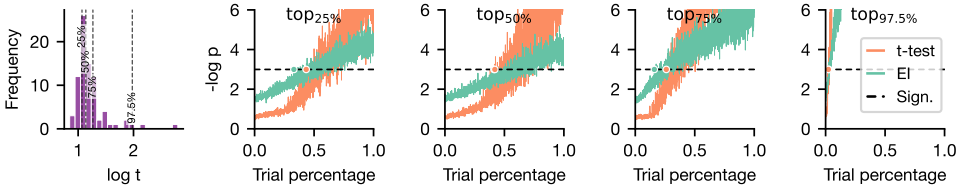
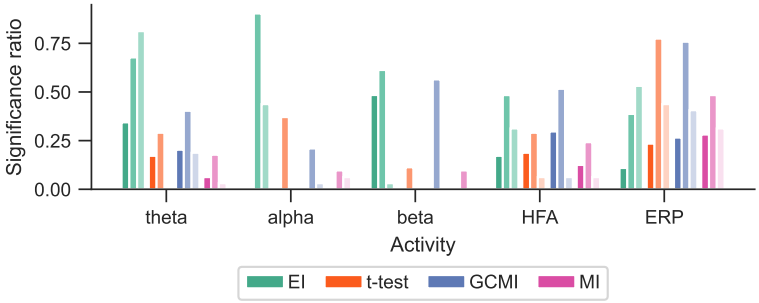
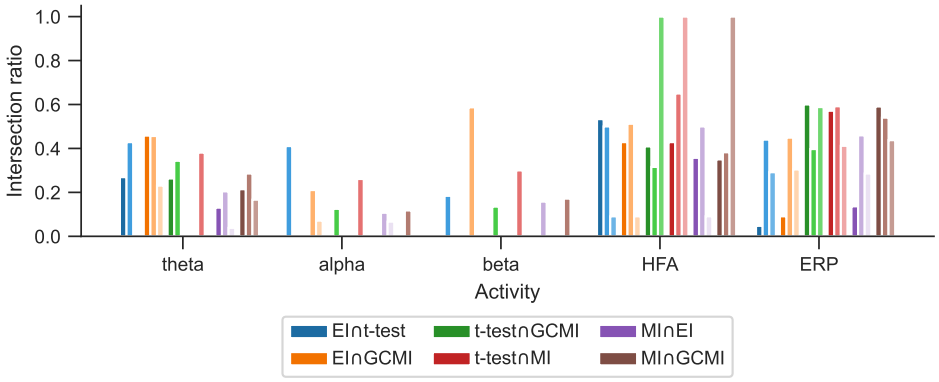


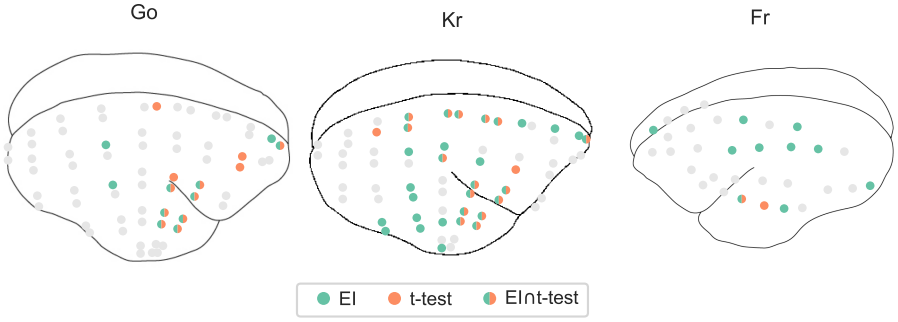
Figure 5: Performance in the case of a limited amount of data. The number of trials was reduced by randomly selecting a varying number of trials for each condition. This was followed by repeatedly applying the measures (50 times for each percentage). The dashed line represents the significance threshold (p=0.05). Note that the y-range is limited to 6 (corresponding to a p-value of $2.5e{-}3$). 100 % of the trials was around 759.50 ± 360.85 trials for deviant responses and 715.14 ± 388.12 trials for standards for each channel. The dots on the significance line indicate the trial percentage when each method exceeded the significance level.



(a) Significance ratios across activities for the three monkeys.

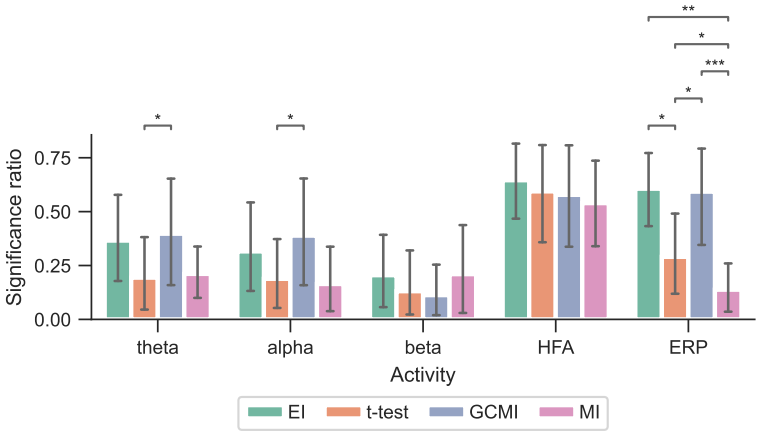


(b) Intersection of significant channels across methods.

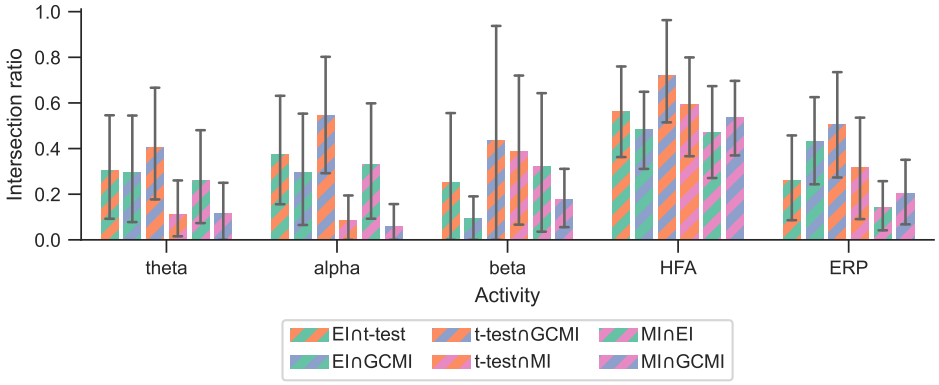


(c) Location of the significant channels based on HFA across the marmosets.

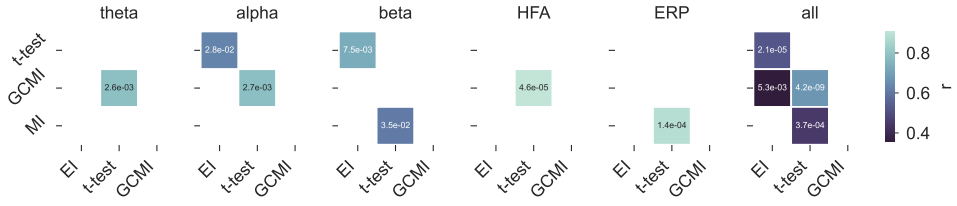
Figure 6: Performance of the different methods across the three marmoset monkeys ”Go”, ”Kr” and ”Fr”. **a**: Global significance ratio for each marmoset. **b**: Intersection of the significant channels for each method combination. Each number is shown relative to the total number of significant channels for each method combination. **c**: Location of the significant channels for EI and t-test for HFA (monkey ”Go” has 64, ”Kr” 62, and ”Fr” exhibits 32 channels).



(a) Significance ratios across activities and methods.



(b) Intersection of significant channels across methods.



(c) Correlation of methods' performances across subjects.

Figure 7: Performance of the different methods across subjects for the vWM task. **a**: Performance of the different methods. The error bars indicate the 95% CIs across subjects. **b**: Significance ratio across iEEG activity types. Statistical significance is indicated with * $p \leq 5e{-}2$, ** $p \leq 1e{-}2$, *** $p \leq 1e{-}3$ and **** $p \leq 1e{-}4$. **c**: Correlation matrices comparing the subject-specific significance ratios. The respective p-value is annotated in each square.

role. By masking a channel's significance, it is necessary to reduce the time course of the test statistics to one concrete value. Additionally, the complete time series of the respective statistic also needs to be corrected for multiple comparisons across channels and trials (i.e., correction of a 2D array that covers samples and channels versus a 1D array that only includes channels).

Adopting the sample-based approach mentioned above is not beneficial for EI. The measure of EI is a compression-based approach and thus compresses the entire time series while exploiting structures along the time axis. This mechanism is a key feature of this approach and is possibly responsible for the higher significance ratios across activities (Fig. 3a, 4, 6a, and 7a). In using the trials for each sample, this feature would be removed. Especially for responses that are close in magnitude (Fig. 1), this mechanism proves beneficial. By exploiting complementary structures within both responses, subtle information-grounded differences that extend beyond the width of one sample can be identified in such scenarios. Additionally, the binning parameter or resolution along the time axis plays a minor importance, while it can have a great effect on the computation time of GCMI or cluster-based t-test.

Especially for beta and HFA activity, EI appears to be a useful tool for detecting active channels (Fig.3a). For all data sets, EI exhibited the highest significance ratios compared to the t-test, MI, and GCMI. HFA presumably carries stimulus mismatch or prediction error signals (24). In that regard, detecting a high number of channels in HFA discriminating between standard and deviating sounds in regions such as the hippocampus appears to be especially interesting (Fig. 4 and 7a).

For the ERP activity, the t-test showed the greatest significance ratios for the optimum-1 and roving oddball paradigm. One possible reason for this could be the generally low variance (or standard error) in EPRs compared to the other electrophysiological activity types. This low variance leads to statistically strong differences in ERP amplitudes resulting from the different stimulus conditions. Further, the number of samples that significantly differ in amplitude after correcting for multiple comparisons can be low for a channel (i.e., the whole trial versus only a few sample points). EI, on the other hand, considers the full-time course and only assesses channels as significant when the information content of the entire mean responses differs. However, the high performance of the t-test is not observable for the vWM task. Here, both EI and GCMI detected a greater number of responsive channels, while also showing a relatively high intersection ratio between t-tests, respectively.

Overall, the sample-based approaches of t-test and GCMI correlated the most (Fig. 3c). Besides that, there was no clear outlier apparent for both data sets. The same conclusion can be drawn considering the significance ratios across individual brain regions (Fig. 4). This can be seen as a validation of the EI measure: While showing higher significance ratios in some ROIs (e.g., hippocampus or pre-central sulcus), it performed similarly well as the other methods across brain areas and did not indicate implausible results.

Compared to EI, t-test, MI, and GCMI, the NFT approach had an inferior performance. However, when considering individual ROIs, it performed equally well for regions close to or belonging to the temporal cortex. Notably, NFT is based on a different procedure. It exploits the static presentation rates of stimuli visible in the spectral decomposition of the entire sequence. Fig. 2 shows such a "tagging" effect, where the brain synchronizes to the exact frequency of the stimuli presentation. The highest SNR or power was contained at 1.92 Hz which was the presentation rate of the standard or deviant tones. In addition, it also showed a peak at half this frequency, which is the combination of either standard-to-standard or deviant-to-deviant tones. This implies that the brain discriminates between standard and deviant tones. One prerequisite here is that the presentation rate is static. Any temporal jitter during the presentation critically disturbs this "tagging" effect. A possible reason for the relatively low performance could be related to this unique feature of NFT, together with the employed multidimensional oddball paradigm. It may perform better in paradigms with fewer deviant types than in the present task containing eight types of deviant tones.

When it comes to performance, MI was close to NFT. It detected fewer channels than the other methods, especially for alpha, beta, and HFA activity. It is important to notice that for comparison reasons, we implemented MI in the same way as EI. That is, the entire time series of a mean response was considered. It is also possible to implement it in the sample-based fashion of GCMI, where it has been shown to operate similarly well (3).

Furthermore, the number of bins during the binning procedure across the data sets was held constant.

100

We chose 128 bins for EI and four bins for MI (3, 5, 10). Adapting this parameter might be needed when the relevant trials are of increasing length, i.e., for longer-lasting trials, it might be necessary to reduce the number of bins for the EI measure. The main motivation for this step is to simplify the signal to an extent that ensures that the compressor keeps operating effectively. However, both EI and MI should be robust to this parameter choice. For example, EI showed similar results when varying the number of bins from 128 to 64, but showed a decreased performance when using only 32 bins (6, 7, 10). An alternative to adapting the number of bins is to segment the time series into shorter segments, which would then yield segments of EI values along the trials, or to down-sample the signal along the time axis.

Lastly, in the case of limited data availability, the two best-performing methods showed relatively similar performance (Fig. 5). For lower t-values (below the 75-percentile, both methods needed around 50 % of the maximally available trials to identify a channel as event-responsive. However, EI appeared to require slightly fewer trials than the t-test. Notice, that for channels with relatively distinct responses to standards and deviants (above the 97.5-percentile), both measures only required around 3 % (ca. 22 trials) of the available trials to identify the respective channel as significant.

Altogether, the examined neurophysiological data sets stem from two species (human and non-human primates) and employ both passive and active paradigms as well as auditory or visual stimuli. Note that in contrast to the optimum-1 and roving oddball paradigms, in the vWM study subjects were instructed to solve a memory task where they were exposed to visual stimuli. Moreover, as opposed to comparing standard with deviant tones responses, for the vWM paradigm, the probe period was compared to the baseline period based on one cortical region. Regardless of these differences, EI performed robustly across all data sets (Fig. 3a, 6a, and 7a).

Conclusion

Taken together, our findings demonstrate that EI stands out as a highly advantageous procedure to discriminate differences in neurophysiological recordings of evoked responses. Among the methods considered, EI competed well in detecting iEEG channels sensitive to deviating sound types across diverse types of electrophysiological responses. Especially for beta and HFA activity, EI detected a higher number of sensitive channels in comparison to the other procedures. Future studies could explore this further by extending EI through a hybrid approach, i.e., using Shannon's information theory besides AIT. Another possibility is to focus on the compression method. Modern compressors such as LZMA, Brotli, or neural network compressors might be able to increase its performance. In sum, our proposed information-based encoding measure proved to be favorable for assessing where in the brain neural responses differ across experimental conditions.

Author Contributions

JF, KG, AKS, TE, and AB contributed to the conception and design of the study. AB and AL carried out the experiment and collected the data. JF, AB, KG, TE, and AKS contributed to the interpretation of the results. JF wrote the manuscript with inputs from AB, KG, TE, and AKS. All authors revised the manuscript. All authors read and approved the final manuscript.

Funding

This work was partly supported by the Research Council of Norway (RCN) through its Centres of Excellence scheme project number 262762, RCN project number 240389, and RCN project number 314925.

Ethics Approval and Consent to Participate

This study was approved by the Research Ethics Committee of El Cruce Hospital, Argentina, the Regional Committees for Medical and Health Research Ethics, Region North Norway, and the University of Cal-

ifornia, Berkeley. Patients gave written informed consent before participation. Electrode placement for the human iEEG recordings was solely determined based on clinical considerations

Acknowledgments

We thank the patients for kindly participating in the studies. We want to express our gratitude to the EEG technicians at El Cruce Hospital and Oslo University Hospital-Rikshospitalet for their support. We thank Pål G. Larsson, Jugoslav Ivanovic, and Ludovic Bellier for their collaboration. Further, we would like to thank Misako Komatsu, Kana Takaura, and Naotaka Fujii for publicly sharing their data.

Data Availability Statement

Custom analysis codes written in Matlab are available at osf.io/tnvc4. Neurophysiological data for the roving oddball paradigm is publicly available at neurotycho.org/auditory-oddball-task. Due to the confidential nature of the data, the patients' datasets analyzed for the optimum-1 paradigm study are not publicly available. The ethical approval conditions do not permit public archiving of study data. Readers seeking access to the data should contact the AB, Department of Psychology, University of Oslo; the Research Ethics Committee of El Cruce Hospital, Argentina; and the Regional Committees for Medical and Health Research Ethics, Region North Norway. Requests must meet the following specific conditions to obtain the data: a collaboration agreement, a data-sharing agreement, and a formal ethical approval. Data for the vWM task can be made available from AL upon reasonable request.

Competing interests

The authors declare no conflict of interest.

References

1. N. M. Timme and C. Lapish, "A tutorial for information theory in neuroscience," *eNeuro*, vol. 5, no. 3, 2018.
2. M. Li and P. Vitányi, *An Introduction to Kolmogorov Complexity and Its Applications*, 3rd ed., ser. Texts in Computer Science. Springer New York, 2008.
3. R. A. Ince, B. L. Giordano, C. Kayser, G. A. Rousselet, J. Gross, and P. G. Schyns, "A statistical framework for neuroimaging data analysis based on mutual information estimated via a gaussian copula," *Human Brain Mapping*, vol. 38, no. 3, pp. 1541–1573, 2017.
4. E. Piasini and S. Panzeri, "Information theory in neuroscience," *Entropy*, vol. 21, no. 1, 2019. [Online]. Available: https://www.mdpi.com/1099-4300/21/1/62
5. J. Fuhrer, K. Glette, J. Ivanovic, P. G. Larsson, T. Bekinschtein, S. Kochen, R. T. Knight, J. Tørresen, A.-K. Solbakk, T. Endestad, and A. Blenkmann, "Direct brain recordings reveal continuous encoding of structure in random stimuli," *bioRxiv*, 2021.
6. J. D. Sitt, J.-R. King, I. El Karoui, B. Rohaut, F. Faugeras, A. Gramfort, L. Cohen, M. Sigman, S. Dehaene, and L. Naccache, "Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state," *Brain*, vol. 137, no. 8, pp. 2258–2270, 06 2014.
7. A. Canales-Johnson, A. J. Billig, F. Olivares, A. Gonzalez, M. d. C. Garcia, W. Silva, E. Vaucheret, C. Ciraolo, E. Mikulan, A. Ibanez, D. Huepe, V. Noreika, S. Chennu, and T. A. Bekinschtein, "Dissociable Neural Information Dynamics of Perceptual Integration and Differentiation during Bistable Perception," *Cerebral Cortex*, vol. 30, no. 8, pp. 4563–4580, 03 2020.
8. M. Schartner, A. Seth, Q. Noirhomme, M. Boly, M.-A. Bruno, S. Laureys, and A. Barrett, "Complexity of multi-dimensional spontaneous eeg decreases during propofol induced general anaesthesia," *PLOS ONE*, vol. 10, no. 8, pp. 1–21, 08 2015.
9. M. Schartner, A. Pigorini, S. A. Gibbs, G. Arnulfo, S. Sarasso, L. Barnett, L. Nobili, M. Massimini, A. K. Seth, and A. B. Barrett, "Global and local complexity of intracranial EEG decreases during NREM sleep," *Neuroscience of Consciousness*, vol. 2017, no. 1, 01 2017.
10. J. Fuhrer, A. Blenkmann, T. Endestad, A.-K. Solbakk, and K. Glette, "Complexity-based encoded information quantification in neurophysiological recordings," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2022, pp. 2319–2323.
11. A. O. Blenkmann, S. Collavini, J. Lubell, A. Llorens, I. Funderud, J. Ivanovic, P. G. Larsson, T. R. Meling, T. Bekinschtein, S. Kochen, T. Endestad, R. T. Knight, and A.-K. Solbakk, "Auditory deviance detection in the human insula: An intracranial eeg study," *Cortex*, vol. 121, pp. 189 – 200, 2019.

12. A. Llorens, L. Bellier, A. Blenkmann, J. Ivanovic, P. Larsson, J. Lin, T. Endestad, A.-K. Solbakk, and R. Knight, "Decision and response monitoring during working memory are sequentially represented in the human insula," *bioRxiv*, 2022.
13. M. Komatsu, K. Takaura, and N. Fujii, "Mismatch negativity in common marmosets: Whole-cortical recordings with multi-channel electrocorticograms," *Scientific Reports*, vol. 5, no. 1, p. 15006, 10 2015.
14. A. Canales-Johnson, A. F. Teixeira Borges, M. Komatsu, N. Fujii, J. J. Fahrenfort, K. J. Miller, and V. Noreika, "Broadband dynamics rather than frequency-specific rhythms underlie prediction error in the primate auditory cortex," *Journal of Neuroscience*, vol. 41, no. 45, pp. 9374–9391, 2021.
15. A. M. Norcia, L. G. Appelbaum, J. M. Ales, B. R. Cottereau, and B. Rossion, "The steady-state visual evoked potential in vision research: A review," *Journal of Vision*, vol. 15, no. 6, pp. 4–4, 05 2015.
16. T. W. Picton, M. S. John, A. Dimitrijevic, and D. Purcell, "Human auditory steady-state responses: Respuestas auditivas de estado estable en humanos," *International Journal of Audiology*, vol. 42, no. 4, pp. 177–219, 2003.
17. R. Oostenveld, P. Fries, E. Maris, and J.-M. Schoffelen, "Fieldtrip: Open source software for advanced analysis of meg, eeg, and invasive electrophysiological data," *Computational intelligence and neuroscience*, vol. 2011, p. 156869, 01 2011.
18. B. L. Foster, B. J. He, C. J. Honey, K. Jerbi, A. Maier, and Y. B. Saalmann, "Spontaneous neural dynamics and multi-scale network organization," *Frontiers in Systems Neuroscience*, vol. 10, 2016.
19. M. Li, X. Chen, X. Li, B. Ma, and P. Vitanyi, "The similarity metric," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3250–3264, 2004.
20. A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, and M. S. Hämäläinen, "MEG and EEG data analysis with MNE-Python," *Frontiers in Neuroscience*, vol. 7, no. 267, pp. 1–13, 2013.
21. T. Meigen and M. Bach, "On the statistical significance of electrophysiological steady-state responses," *Documenta Ophthalmologica*, vol. 98, no. 3, pp. 207–232, Jul 1999.
22. R. Q. Quiroga, Z. Nadasdy, and Y. Ben-Shaul, "Unsupervised Spike Detection and Sorting with Wavelets and Superparamagnetic Clustering," *Neural Computation*, vol. 16, no. 8, pp. 1661–1687, 08 2004.
23. D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.
24. A. M. Bastos, J. Vezoli, C. Bosman, J.-M. Schoffelen, R. Oostenveld, J. Dowdall, P. De Weerd, H. Kennedy, and P. Fries, "Visual areas exert feedforward and feedback influences through distinct frequency channels," *Neuron*, vol. 85, no. 2, pp. 390 – 401, 2015.