Peer Community In Registered Reports
Stage 1 Proposal **Submission #3**

Detecting differences in conscious contents using EEG complexity measures

(Anonymous submission)

Abstract

Measuring consciousness has been a longstanding problem. Even though behavioral responses are commonly used, converging evidence indicates that behavioral responsiveness and behavioral reports about consciousness dissociate from consciousness per se. Measures of complexity applied to brain activity, such as Lempel-Ziv complexity (LZc) and the perturbational complexity index (PCI), have been shown to discriminate between levels of consciousness, but less of this work has been done in the context of conscious *content*. To address many of the limitations of previous work, in this study we measure participants' neurophysiological (EEG), subjective, *and* behavioral responses in states of normal wakefulness to visual *and* auditory stimuli that vary in granularity of subjective characteristics, such as meaningfulness. Two novel aspects of our study are that some of the visual and auditory stimuli are manipulated such that on most initial trials they are unrecognizable, but on some subsequent trials, they become recognizable. This allows us to measure changes in EEG complexity that correspond to differences in phenomenology alone while completely controlling for stimulus complexity. In addition, we are assessing if any of five dimensions of subjective ratings correlate with any differences in EEG complexity. This study advances our understanding of consciousness by clarifying the relationship between stimulus complexity and measures of brain complexity and phenomenology.

Cover Letter

To our knowledge, this study will be the first of its kind to do *all* of the following: 1) compute measures of brain complexity on electroencephalography signals; 2) compute measures of brain complexity of both functional differentiation and of the joint presence of functional differentiation and functional integration; 3) compute measures of brain complexity in response to both visual and auditory stimuli; 4) compute measures of brain complexity in eyes-open and eyes-closed conditions (for the auditory paradigm); 5) compare measures of brain complexity at different levels of granularity of stimuli; 6) correlate differences in subjective ratings of experiences to stimuli with differences in measures of brain complexity; and 7) include at least 25 ~~30~~ participants. All necessary facilities and equipment are in place for the proposed research. Funding is being provided through the Funding Consciousness Research with Registered Reports initiative (https://www.cos.io/consciousness), offered by the Center for Open Science, Templeton World Charity Foundation, and the Association for the Scientific Study of Consciousness. The research protocol has been submitted to our Institutional Review Board, and ethical approval is currently pending. Upon Stage 1 in principal acceptance (IPA), data collection is estimated to take 4-5 months, data analysis 2-3 months, and final manuscript preparation 2-3 months. All authors agree to share their raw data, any digital study materials, and analysis code as appropriate. Following Stage 1 IPA, all authors agree to register their approved protocol on the Open Science Framework. If the authors later withdraw their paper, they agree to a short summary of the pre-registered study being published under the article type Withdrawn Registrations.

Estimated timeline.

<table>
<tr><td rowspan="2">IPA</td><td>Month 1</td><td>Month 2</td><td>Month 3</td><td>Month 4</td><td>Month 5</td><td>Month 6</td><td>Month 7</td><td>Month 8</td><td>Month 9</td><td>Month 10</td><td>Month 11</td></tr>
<tr><td colspan="5">Data collection</td><td colspan="3">Data analysis</td><td colspan="3">Final manuscript preparation</td></tr>
</table>

1. Introduction

Measuring consciousness has been a longstanding problem (Gosseries et al., 2014; Overgaard, 2015; Sandberg et al., 2010; Seth et al., 2008), as has been defining it (Block, 1995; Overgaard, 2015; Tassi & Muzet, 2001; Velmans, 2009). Three common ways to define consciousness are: 1) awareness of self and environment (Giacino et al., 2014; James, 1894); 2) "what it is like" to be something ~~have an experience~~ (Nagel, 1974); and 3) the difference between wakefulness and dreamless sleep (Velmans, 2009) ~~the capacity to have any experience (Sarasso et al., 2021)~~. "What-it-is-like-ness" can be thought of in terms of ~~has also been referred to as local states of consciousness, or~~ conscious content; and the other two definitions can be thought of in terms of ~~whereas the capacity to have any experience has been referred to as global states of consciousness, or~~ conscious level[1] (Bayne et al., 2016; Seth & Bayne, 2022).

Two ways to measure consciousness are based on behavior and based on brain activity (Seth et al., 2008). Converging evidence, however, indicates that behavioral responsiveness and behavioral reports about consciousness dissociate from consciousness per se (Frässle et al., 2014; Koch et al., 2016; Owen et al., 2006; Pitts et al., 2014; Sanders et al., 2012; Sarasso et al., 2015). For this reason, the most robust measures of consciousness will be brain-based rather than behavioral. Measures of consciousness should also be generic rather than context-specific, so that they can be deployed across the wide range of circumstances in which measuring consciousness is of interest (Koculak & Wierzchón, 2022; Orlowski & Bola, 2023). Such circumstances include diagnosing disorders of consciousness (Gosseries et al., 2014; Owen et al., 2006), monitoring anesthesia depth (Kissin, 2000; Ní Mhuircheartaigh et al., 2013; Rampersad & Mulroy, 2005; Warnaby et al., 2017), designing no-report paradigms (Pitts et al., 2014; Tsuchiya et al., 2015), and assessing consciousness in infants and fetuses (Bayne et al., 2023; Frohlich et al. 2023), and. In this study, we focus on *brain*-based measures for conscious level and conscious content.

Measures of brain complexity have been shown to discriminate between levels of consciousness (Frohlich et al., 2021), but less of this work has been done in the context of conscious content. Some studies have shown that such measures can discriminate between participants' experiences of meaningful and non-meaningful visual stimuli (Boly et al., 2015; Mensen et al., 2017, 2018), but these studies: 1) have not shown consistent results for auditory stimuli (Canales-Johnson et al., 2020; Bola et al., 2018; Orlowski & Bola, 2023), nor for certain complexity measures (Bola et al., 2018; Orlowski & Bola, 2023) and for eyes-open vs. eyes-closed conditions with psychedelics (Farnes et al., 2020; Mediano et al., 2024). Furthermore, these studies: 2) have not evaluated a key complexity measure of conscious level; 3) have not investigated fine-grained differences in conscious contents; and 4) have mostly been based on small sample sizes. In addition, it is an open question as to exactly which aspects of phenomenology these measures may reflect (Bola et al., 2018; Carhart-Harris, 2018; Mediano et al, 2020; Murray et al., 2024; Schartner et al., 2017a; Timmermann et al., 2019).

[1] Although "the difference between wakefulness and dreamless sleep" implies a binary distinction, conscious "level" also typically implies graded delineations (Bayne et al., 2016) ~~In the context of consciousness, "level" can also refer to graded delineations within "capacity", but we use the two terms interchangeably for both graded and binary descriptions of consciousness.~~

This study addresses these issues and advances our understanding of consciousness by clarifying the relationship between brain complexity and phenomenology.

In section one of this paper, we review the aforementioned literature relevant to measuring conscious level, which includes two complexity measures in particular: Lempel-Ziv complexity and the perturbational complexity index. We also review the smaller body of research that pertains to complexity measures and conscious *content*, including studies that have investigated changes in brain complexity induced by psychedelics. Section one concludes with a description of our experimental aims and hypotheses. In section two, we detail our experimental methods.

1.1. Measuring conscious level

Many brain-based measures of conscious level have been proposed, such as approximate entropy (Pincus et al., 1991), spectral entropy (Johnson & Shore, 1984), and the bispectral index (Kissin, 2000; Sigl & Chamoun, 1994). Each of these measures can be understood as quantifying the information content of brain activity. Other proposed measures of conscious level include late ERPs (Plourde & Picton, 1991), measures of effective connectivity (Boly et al., 2011; Rosanova et al., 2012), and Granger causality of electrophysiological (Engel & Singer, 2001) or metabolic (Vanhaudenhuyse et al., 2010) signals. Each of these measures can be understood as quantifying the spatial extent or synchronization of brain activity. None of these measures, however, have proven reliable within subjects and across the many different conditions of consciousness (Casali et al., 2013;).

A more promising approach to measuring conscious level has been to measure both of these dimensions of brain activity. This approach can be traced to the first paper explicitly linking consciousness with complexity by Tononi & Edelman (1998) (Sarasso et al., 2021). For Tononi & Edelman (1998), neural complexity (which was also the name of their proposed measure) consists of both functional differentiation *and* functional integration. Functional differentiation refers to the repertoire of different brain states, whereas functional integration refers to how unified different parts of the brain are interacting. A key claim of Tononi & Edelman (1998) is that the kind of complexity that matters for consciousness is the coexistence of a high degree of both functional differentiation and integration.

1.1.1. Spontaneous Lempel-Ziv complexity and conscious level

A generic complexity measure that has outperformed many of the aforementioned measures of conscious level is Lempel-Ziv complexity (LZc) (Bai et al., 2015; Frohlich et al., 2021; Schartner et al., 2015; Schartner et al., 2017b; Zhang et al., 2001). In general, LZc quantifies how redundant or compressible any binary sequence is by estimating the rate at which distinct substrings are encountered upon scanning the sequence (Lempel & Ziv, 1976; Ziv & Lempel, 1977). Specifically, the algorithm: 1) scans the sequence from left to right; 2) finds the shortest substring that hasn't yet been encountered; 3) adds this substring to a dictionary of previously encountered substrings; 4) advances the sequence just beyond the last substring; and 5) repeats the previous steps until the end of the sequence is reached. The value of LZc corresponds to the total number of substrings added to the dictionary. LZc can be applied to electroencephalography (EEG) signals[2] recorded in a passive manner ("spontaneously") or an active manner ("perturbational"). LZc was first applied to EEG signals to study epilepsy by Radhakrishnan & Gangadhar (1998), and then to study depth of anesthesia by Zhang et al. (2001). Spontaneous LZc[3] can be understood to be estimating the repertoire of brain states via temporal differentiation, but without explicitly estimating the integration dimension of the signal (Sarasso et al., 2021; Schartner et al., 2015).

Spontaneous approaches may be more limited than perturbational approaches for estimating differentiation, because they only capture brain states that happen to be cycled through during the observational periods rather than states that could be cycled through if the brain were to be perturbed accordingly (Sarasso et al., 2021; Massimini et al., 2009). Furthermore, to the extent that spontaneous approaches reflect some degree of integration, they likely overestimate it due to not accounting for spurious correlations (Sarasso et al., 2021; Massimini et al., 2009; Ort et al., 2023).[4] Perturbational approaches may more accurately estimate integration by accounting for spurious correlations similarly to frameworks that employ interventions for understanding causation (Holland, 1986; Pearl & Verma, 1995). Although spontaneous LZc has been shown to track the depth of anesthesia, there is evidence of a potential dissociation between LZc and consciousness in rats (Frohlich et al., 2021; Pal et al., 2020).

1.1.2. Perturbational complexity index and conscious level

Another complexity measure that appears to outperform LZc in discriminating conscious level is the perturbational complexity index (PCI) (Frohlich et al., 2021). PCI was developed as a measure of conscious level to gauge the amount of information contained in the brain's spatiotemporal response to a transcranial magnetic stimulation (TMS)-induced cortical perturbation using EEG. PCI thus employs a perturbational approach, and it explicitly gauges both differentiation and integration. It gauges differentiation by considering the information contained in the response, and it gauges integration by considering the spatiotemporal characteristics of the response. PCI has been shown to discriminate between consciousness and unconsciousness within healthy subjects, as well as across levels of consciousness in brain-injured patients with unprecedented sensitivity and specificity (Casali et al., 2013; Sinitsyn et al., 2020).

There are two formulations of PCI, and they utilize different computational methods to estimate differentiation and integration. The original formulation, which can be referred to as PCIlz since it uses LZc in its algorithm, was introduced by Casali et al. (2013). The first step of the PCIlz algorithm is to binarize the continuous EEG signal at the source level[5] using a voltage-amplitude threshold[6] corresponding to significant differences in cortical activations between the post-TMS

[2] ~~LZc was first applied to EEG signals to study epilepsy by Radhakrishnan & Gangadhar (1998), and then to study depth of anesthesia by Zhang et al. (2001).~~

[3] From here onward, LZc refers to spontaneous LZc unless otherwise noted or implied.

[4] ~~Perturbational approaches may more accurately estimate integration by accounting for spurious correlations similarly to frameworks that employ interventions for understanding causation (Holland, 1986; Pearl & Verma, 1995).~~

[5] ~~Source localization for PCIlz requires using high-density (HD) EEG~~.

[6] ~~The voltage-amplitude threshold for PCIlz is determined from a non-parametric statistical approach~~.

response and pre-TMS baseline periods. Source localization for PCIlz requires using high-density (HD) EEG, and the voltage-amplitude threshold for PCIlz is determined from a non-parametric statistical approach. Binarization based on significant sources ~~This step~~ can be understood as an estimation of integration, because the resulting matrix reflects only activity from those parts of the brain that have responded in a unified manner to the TMS perturbation. The next step of the PCIlz algorithm is to compute LZc on the resulting binarized spatiotemporal matrix of significant source activations,[7] which yields an estimation of differentiation. The final PCIlz measure is also normalized by source entropy.

In general, PCIlz is low either when integration is low, because reduced interactions among cortical areas yield a spatially restricted activation matrix, or when differentiation is low, because stereotypical responses yield a more compressible activation matrix (Casali et al., 2013). PCIlz is high only when a large number of brain areas respond to the TMS perturbation together but in a differentiated manner, yielding a spatiotemporal activation matrix that is not easily compressible (Ibid.). In healthy subjects, the highest values of PCIlz correspond to states of wakefulness, and the lowest values correspond to states of non-rapid-eye movement (NREM) sleep and anesthesia. For example, propofol, which is an anesthetic that potentiates gamma-amino butyric acid (GABA) receptor activity, yields a spatially restricted activation matrix consistent with loss of integration; whereas xenon, which is an anesthetic that potentiates the conductance of 2PK+ channels and antagonizes N-methyl-D-aspartate (NMDA) receptors, yields a widespread but stereotypical (and thus compressible) activation matrix consistent with loss of differentiation (Sarasso et al., 2015).

The second formulation of PCI is known as PCIst and was introduced by Comolatti et al. (2019). PCIst utilizes principal component decomposition and state-transition quantification, and can thus be computed on any evoked brain signal.[8] By computing principal components, PCIst mitigates the need for source-localization and thus HD-EEG. PCIst is also much faster to compute than PCIlz. The steps of the PCIst algorithm are to: 1) decompose the TMS-evoked potentials into principal components; 2) calculate voltage-amplitude distance matrices for the baseline and response samples (for each principal component); 3) threshold these matrices to 4) yield temporal transition matrices; and 5) sum across the maximized differences between the number of state transitions in the response and in the baseline (for all selected principal components). The result is an (unnormalized) scalar value that is significantly correlated with PCIlz but whose magnitude also scales with the number of EEG channels. PCIst is high when there are multiple spatially distributed linearly independent components, each of which contribute significant numbers of state transitions to the brain's overall response to the perturbation (Ibid.). For this formulation of PCI, principal component decomposition and state-transition quantification roughly correspond to estimating integration and differentiation, respectively.

~~[7] The final PCIlz measure is also normalized by source entropy.~~

~~[8] By computing principal components, PCIst mitigates the need for source-localization and thus HD-EEG. PCIst is also much faster to compute than PCIlz.~~

To the extent that PCI outperforms LZc in discriminating conscious level, one possible explanation is that the kind of complexity that matters for consciousness is indeed the coexistence of a high degree of both functional differentiation and integration (Tononi & Edelman, 1998). In this context, PCI may more effectively estimate *both* of these dimensions of brain activity, because spontaneous LZc by itself doesn't explicitly or accurately estimate the integration dimension, nor does it account for as many possible brain states (differentiation) as perturbational approaches. Another possible explanation is that PCI via TMS-perturbation results in a much higher signal-to-noise ratio compared to LZc via spontaneous activity.

1.2. Measuring conscious content

Compared to measures of conscious level, few brain-based measures of conscious content have been proposed.[9] In a coarse sense, both blood-oxygenation-level-dependent (BOLD) responses from functional magnetic resonance imaging (fMRI) and ERPs from EEG recordings have long been capable of distinguishing brain activity correlated with different perceptual contents (assuming corroborating behavioral reports), such as between faces and other objects (Bentin et al., 1996; Bötzel et al., 1989; Kanwisher et al., 1997). Furthermore, BOLD responses have even been used to discriminate between different putatively conscious mental imagery tasks in patients previously diagnosed as behaviorally unconscious (Boly et al. 2007; Monti et al., 2010; Owen et al., 2006). However, BOLD responses and ERPs are not ~~general purpose~~ conventional mathematical measures ~~on their own~~.[10]

1.2.1. Complexity measures and conscious content

Regarding measures of complexity, there have also been fewer studies investigating their capabilities in the context of conscious content compared to conscious level, perhaps in part because the magnitude of effects may be smaller (Koculak & Wierzchón, 2022). The first such study was by Boly et al. (2015), and they computed LZc, neural complexity (Tononi et al., 1994), and a measure of mutual information[11]. All three measures were computed on fMRI BOLD responses of six participants to three types of video clips varying in meaningfulness: the original clips, clips scrambled in time ("scrambled"), and clips scrambled in space (to produce imagery akin to "TV noise"). In this study, LZc was computed in a perturbational manner, using the onset

[9] Three canonical *non*-brain-based measures for conscious content include: 1) physiological measures such as eye movements and pupil size, which can be used to discriminate between perceptual contents such as those experienced during binocular rivalry (Fox et al., 1975; Frässle et al., 2014; Leopold et al., 1995; Logothetis & Schall, 1990; Tsuchiya et al., 2015); 2) objective measures such as those using forced-choice discriminations (Dulany, 1997; Eriksen, 1960; Pessoa et al., 2006; Smyth & Shanks, 2008) and signal detection theory (Green & Swets, 1966); and 3) subjective measures such confidence ratings (Cheesman & Merikle, 1986; Dienes et al., 1995) and the perceptual awareness scale (Ramsøy & Overgaard, 2004; Sandberg & Overgaard, 2015).

[10] Although ERPs are not general purpose measures, per se, and although the P3b ERP, which was once suggested as a signature of conscious processing (Bekinschtein et al., 2009) has since been deemed not to be a marker of consciousness (Faugeras et al., 2011; Fischer et al., 2010; Holler et al., 2011; Koch et al., 2016; Kotchoubey, 2005; Sitt et al., 2014; Tzovara et al., 2015), the perceptual ~~visual~~ awareness negativity ERP may reliably correlate with conscious perception (Dembski et al., 2021; Koch et al., 2016; Pitts et al., 2014; Railo et al., 2011).

[11] The third measure computed by Boly et al. (2015) assessed the difference between the mutual information of the whole system and that of its parts (Oizumi et al., 2010; Oizumi et al., 2012).

of the video clips relative to a black screen baseline as the time-locking event.[12] They found that all three measures were highest for the original clips, lowest for “TV noise”, and intermediate for the scrambled clips (accounting for stimulus complexity), supporting the authors’ hypothesis that neurophysiological differentiation reflects phenomenological differentiation and the overall meaningfulness of the stimuli. It is worth noting however that while fMRI gives superior spatial resolution compared to EEG, analyzing BOLD responses at the time scale of seconds necessarily minimizes the estimate of differentiation by reducing the repertoire of possible states (Sarasso et al., 2021). To address this issue, we will use EEG in this study.

In a set of studies by Mensen et al. (2017, 2018), they developed a novel measure of neurophysiological differentiation (DA)[13] and applied it to EEG signals in response to images and video clips varying in meaningfulness. In this context, meaningfulness corresponds with any phenomenally distinguishable difference, which must be supported by some neurophysiological difference (Ibid.). Thus, the relationship between meaningfulness and consciousness is that meaningfulness is a way to characterize differences between experiences, which is to say, between conscious contents. In Mensen et al. (2017), they measured DA for nine participants in response to four types of images: natural, random noise, random three-dimensional spheres, and phase-scrambled. In this study, they controlled for stimulus predictability and novelty, and also collected subjective ratings for how distinct (phenomenally distinguishable) the images were within a category. They found that DA was high for meaningful images and low for meaningless ones, even at the individual level. They also accounted for the differentiation of the stimuli themselves, and for ERP changes, but they were unable to make definitive statements about which specific categories were significantly different from each other due to post-hoc comparison alpha-level limitations. In this study, we will investigate these finer-grained differences in conscious contents by utilizing Bayesian analysis methods, which allows us to compute posterior distributions for all parameters of interest.

In Mensen et al. (2018), they measured DA for eight participants in response to eight types of video clips varying in meaningfulness: original (habituated and novel), reversed in time, outlined, shuffled, phase-scrambled, phase-scrambled with temporal conservation, and “TV noise”. They also collected subjective ratings for how interesting, meaningful, and understandable the video clips were, and for how many different experiences participants had during the clips. The results from this study extended the findings from Mensen et al. (2017), adapting DA to single trials of spontaneous EEG recordings, but none of these studies so far investigated whether these findings would generalize to other sensory modalities.

[12] Two stimulus designs were used in Boly et al. (2015) that varied in stimulus presentation duration so that LZc could be compared to the black screen baseline in one design (20-second presentation), and to mean BOLD activity to yield systematic BOLD activations and deactivations in the other design (4-minute presentation).
[13] In Mensen et al. (2017), DA was based on differences between stimulus-evoked (time-locked) EEG activity within a stimulus set. In Mensen et al. (2018), DA was based on differences in power spectral density between all states during a single clip, because the video clips used weren’t conducive to time-locked comparisons.

In a final set of studies by Bola et al. (2018) and Orlowski & Bola (2023), they extended the investigation of complexity measures and conscious content to the auditory modality and with larger sample sizes. In Bola et al. (2018), they computed two versions of LZc on spontaneous EEG signals recorded from 19 participants listening to five different speeds of audio books (varying in information rate), plus backwards-played clips (unintelligible) and a resting-state condition. They found for both versions of LZc that the speed of the audio books had no significant effect, whereas LZc was actually greater during the *resting-state* condition (an effect in the opposite direction compared to previous studies with visual stimuli). Furthermore, there was no difference between the original speed (meaningful) and backwards-played audio clips. Interestingly, in a study by Canales-Johnson et al. (2020), they computed various differentiation and integration measures on spontaneous EEG signals recorded in response to bistable auditory tones, and they also found an effect for a version of LZc in the opposite direction than that of a measure designed to assess the integration dimension of the signal.

In Orlowski & Bola (2023), they again computed both versions of LZc on spontaneous EEG signals recorded from 24 participants but while listening to audiobooks *or* watching videos. The stimuli consisted of three versions of audiobooks and videos varying in meaningfulness: original, temporally shuffled ("scrambled"), and noised (plus a resting-state condition). They found for one version of LZc that it was higher for original video clips than it was for noised clips, but that LZc was *lower* for original and scrambled *audio* clips than it was for noised clips.[14] Contrary to previous studies with visual stimuli, they found no difference between original and scrambled clips (nor with auditory stimuli). Having corroborated their previous findings that LZc increases for meaningful visual stimuli but decreases for meaningful auditory stimuli, they concluded that measures of spontaneous EEG signal diversity are not generic indexes of the variability of conscious experience. However, the main limitations of these latter two studies were that they did not collect any subjective reports/ratings, and they did not evaluate any measures that explicitly gauged both differentiation and integration, such as PCI. In this study, we will collect subjective ratings and compute PCI, as well as investigate both visual and auditory modalities.

1.2.2. Psychedelics, complexity measures, and conscious content

A final area of research that has investigated complexity measures in the context of conscious content has been during altered states of consciousness induced by psychedelics and other drugs. In a study by Schartner et al. (2017a), they computed two versions of LZc and two related entropy measures on spontaneous magnetoencephalographic (MEG) signals recorded from participants after receiving psilocybin, lysergic acid diethylamide (LSD), and subanesthetic doses of ketamine[15]. They found that LZc and one of the other entropy measures *increased* in participants (with their eyes open) after receiving all three psychedelics compared to placebo

[14] The second version of LZc computed by Orlowski & Bola (2023) appeared less sensitive than the first in that it did not replicate all of the effects observed for the first version. In this study, we will measure both of these versions of LZc.

[15] Ketamine, although an anesthetic with a different mechanism of action than the "classic" psychedelics (NMDA antagonism vs. 5-HT2A agonism), at *subanesthetic* doses reliably induces dissociative, hallucinogenic, and other subjective effects similar to those of the classic psychedelics (Schartner et al., 2017a).

(accounting for changes in the spectral profile).[16,17] Furthermore, they found strong correlations between the signal diversity measures and many subjective ratings of the experiences, such as intensity.

Subsequent studies have shown similar increases in spontaneous MEG/EEG LZc with N,N-dimethyltryptamine (DMT) (Timmermann et al., 2019), LSD (Mediano et al., 2024), and dose-dependent LSD (Murray et al., 2024). In the study by Mediano et al. (2024), LSD-induced increases in LZc were actually greatest when participants' eyes were closed and not perceiving any visual (nor auditory) stimuli. In the study by Murray et al. (2024), however, greater LZc was neither sufficient to induce subjective effects with low doses of LSD, nor necessary to induce subjective effects with delta-9-tetrahydrocannabinol (THC). Furthermore, another study found no increase in spontaneous EEG LZc with sub-perceptual doses of psilocybin (Cavanna et al., 2022), so it remains an open question as to exactly which aspects of subjective experience (if any), such as intensity or richness, LZc may be indexing (Bola et al., 2018; Carhart-Harris, 2018; Mediano et al, 2020; Murray et al., 2024; Schartner et al., 2017; Timmermann et al., 2019). To address this issue, in this study we will collect ratings for five dimensions of phenomenology to explore if any of these dimensions correlate with changes in complexity.

A final set of studies by Farnes et al. (2020) and Ort et al. (2023) compared spontaneous EEG LZc (and related entropy measures) to PCI in ketamine- and psilocybin-induced altered states of consciousness (respectively). Farnes et al. (2020) found that while LZc and related measures all increased with sub-anesthetic doses of ketamine compared to normal wakefulness and correlated with subjective assessments (both of which were consistent with previous findings), PCI did not. Contrasting with findings from Mediano et al. (2024), Farnes et al. (2020) also found that signal diversity measures were highest in the eyes-open condition. Their suggested explanation for the difference between LZc and PCI was that PCI may reflect the brain's general capacity to sustain consciousness (which was demonstrably preserved), whereas LZc may reflect the complexity of specific conscious content. However, the latter part of this explanation conflicts with the findings from Murray et al. (2024) mentioned in the prior paragraph.

In Ort et al. (2023), they found that while LZc increased with psilocybin compared to placebo (with eyes open but not eyes closed), neither PCIlz nor PCIst changed. While they concluded that PCI may thus not be a measure of specific features of phenomenology, they did find that other aspects of the TMS-evoked responses, such as 10-25 Hz event-related spectral perturbation power, were altered and highly correlated with phenomenological effects such as blissfulness and unity. Their suggested explanation for the difference between LZc and PCI was

---

[16] The increases in LZc and the other entropy measure were significant only for ketamine and LSD (not psilocybin), and they were strongest for ketamine (Schartner et al., 2017a). Potential reasons for these findings were not discussed.

[17] It has also been shown that at *anesthetic* doses, ketamine increases complexity measures that assess randomness (i.e., differentiation), whereas propofol decreases them (Wang et al., 2017). However, ketamine can also increase complexity measures that assess the *balance* between randomness and regularity (akin to levels observed during REM sleep), ~~whereas propofol only decreases them~~ (Ibid.; Sarasso et al., 2015). These ~~latter~~ increases in complexity may explain the fact that anesthetic doses of ketamine also often induce vivid dream-like experiences (Wang et al., 2017; Sarasso et al., 2015).

that the brain was in a state (with psilocybin) that allowed ongoing activity to be more chaotic[18] but which was unchanged in its general ability to maintain complex causal interactions. To address these issues, in this study we will investigate PCI and LZc in the context of conscious content (but not with psychedelics), as well as in eyes-open vs. eyes-closed conditions (in the auditory paradigm).

1.3. Overview of the experiment

The overall aim of this study is to investigate whether the perturbational complexity index (sPCIst) and Lempel-Ziv complexity (sLZc)[19] discriminate between visual- and auditory-evoked[20] differences in conscious contents varying in granularity. We will also explore multiple dimensions of phenomenology that these complexity measures may reflect, such as meaningfulness and intensity. To our knowledge, this study will be the first of its kind to do all of the following: 1) compute complexity measures on EEG signals; 2) compute measures of both functional differentiation and of the joint presence of functional differentiation and functional integration; 3) compute complexity measures in response to both visual and auditory stimuli; 4) compute complexity measures in eyes-open and eyes-closed conditions (for the auditory paradigm); 5) compare complexity measures at different levels of granularity of stimuli; 6) correlate differences in subjective ratings with differences in complexity measures; and 7) include at least 25 ~~30~~ participants. Altogether, this study advances our understanding of consciousness by clarifying the relationship between stimulus complexity and measures of brain complexity and phenomenology ~~conscious content~~, and by addressing several limitations of previous studies.

This study contains two experimental paradigms (visual and auditory) involving a single group of participants. In each paradigm, we will measure the participants' neurophysiological (EEG), subjective, and behavioral responses in states of normal wakefulness to visual and auditory stimuli varying in granularity of subjective characteristics, such as meaningfulness. After a subset of trials, participants will provide subjective ratings for their experience of the stimulus according to the following five dimensions: 1) diversity/richness; 2) unity/integratedness; 3) meaningfulness; 4) intelligibility/understandability; and 5) intensity/vividness. After the same subset of trials, to control for attention and task demands, and allow us to explore complexity on incorrect response trials, participants will also choose which stimulus category/class they

[18] Converging evidence from computer science and neuroscience supports the notion that cognitive and conscious cortical electrodynamics are poised near a boundary between order and chaos (Bertschinger & Natschläger, 2004; Carhart-Harris et al., 2014; Cocchi et al., 2017; Toker et al., 2022). In particular, Toker et al. (2022) demonstrated in computational and electrophysiological data that psychedelic (and unconscious) states reflect transitions of low-frequency cortical electric oscillations towards (and away from) this critical point, as measured by LZc.

[19] To keep as many aspects of the experiment as similar as possible, we compute PCIst and LZc in the same sensory-evoked manner, which we refer to as sPCIst and sLZc, respectively. It is worth noting that Frohlich et al. (2023) has also called for the development of sensory PCI in the context of measuring consciousness in infants and fetuses, where TMS can't be used.

[20] By employing sensory stimuli for PCI and computing LZc in the same evoked manner, we may also be able to gain insight as to whether any difference in discriminatory power between PCI and LZc is due to the higher signal-to-noise ratio of TMS-perturbation.

experienced. For each trial, we will compute sPCIst and sLZc on the evoked-EEG signals recorded in response to the sensory stimuli.

1.3.1. The visual paradigm

In the visual paradigm, we will use three classes of images inspired by Mensen et al. (2017): 1) natural ("natural images"); 2) blurred natural images ("blurred images"); and 3) images of randomly shuffled natural images ("visual noise"). The natural images are composed of two categories: 1) famous people ("famous-people images"); and 2) household objects ("household-objects images"). One novel aspect of this study is that the amount of blurring applied to the natural images will be calibrated such that the blurred natural images are recognizable typically only if a natural version has been seen previously. The blurred images will then be presented on some trials after the natural images so that they become recognizable on those subsequent trials. This will allow us to completely control for low- and high-level stimulus complexity (which is identical between trials) while allowing for differences in phenomenology to manifest between recognizable and unrecognizable visual experiences.

1.3.2. The auditory paradigm

In the auditory paradigm, we will use three classes similar to those of the visual paradigm. For "natural audio", we will translate each natural image into a word and record it being spoken by a native English speaker. Natural audio will be composed of two categories that mirror those of the visual paradigm, plus a third eyes-closed condition: 1) natural images of famous people's names translated into speech ("famous-names audio"); 2) natural images of household objects translated into speech ("household-objects audio"); and 3) natural images of household objects translated into speech with eyes closed ("household-objects audio eyes-closed").[21] In addition to natural audio, there will be two putatively non-meaningful auditory classes: 1) noise-vocoded natural audio ("noise-vocoded audio"); and 2) randomly shuffled natural audio, with eyes open and eyes closed ("auditory noise" and "auditory noise eyes-closed", respectively). As is the case for the blurred images, the noise-vocoded audio clips, which are typically recognizable only if the natural audio has been played previously, will be presented on some trials after the natural audio so that they become recognizable on the subsequent trials. This will again allow us to completely control for stimulus complexity while allowing for differences in phenomenology to manifest between recognizable and unrecognizable auditory experiences.

1.4. Specific aims and hypotheses

1.4.1. Aim 1: granularity of stimuli

Our first aim is to determine if sPCIst and sLZc discriminate between brain responses to visual and auditory stimuli varying in granularity. To do so, we will look for differences in each complexity measure between two levels of granularity of the stimuli. Although there is theoretical justification for a directional relationship between conscious level and PCI, we chose not to impose that directionality on our hypotheses, because there is less theoretical and empirical support for a consistent directional relationship between PCI and LZc and conscious content. Our primary aim is therefore to establish robust, unbiased differences in brain complexity without relying on ground-truth information about subjective experience, and our first exploratory

[21] All auditory conditions are with eyes open unless otherwise noted.

aim focuses on identifying which (if any) dimensions of subjectivity correspond to these complexity differences, given that such relationships remain largely unestablished.

In the visual paradigm, the levels are between: 1) image *classes*, such as between natural and blurred images; and 2) categories *within* the natural image class, such as between famous-people and household-objects images. This aim allows us to determine if sPCIst and sLZc are sensitive (and how sensitive) to changes in sensory-induced brain activity, and to provide within-category evidence not provided by previous studies (Mensen et al., 2017). Our five hypotheses for the visual paradigm of aim 1 are that there will be differences in sPCIst and sLZc between: 1) natural and unrecognizable blurred images; 2) natural images and visual noise; 3) blurred images and visual noise; 4) famous-people and household-objects images, and 5) unrecognizable and recognizable blurred images. All of the hypotheses for both paradigms are presented in tables 1-3.

In the auditory paradigm, the two levels of granularity are between: 1) audio *classes*, such as natural and noise-vocoded audio; and 2) categories *within* the natural audio class, such as between famous-people and household-objects audio. Our five hypotheses for the auditory paradigm of aim 1 are that there will be differences in sPCIst and sLZc between: 1) natural and unrecognizable noise-vocoded audio; 2) natural audio and auditory noise; 3) unrecognizable noise-vocoded audio and auditory noise; 4) famous-names and household-objects audio, and 5) unrecognizable and recognizable noise-vocoded audio.

1.4.2. Aim 2: eyes open vs. eyes closed

Our second aim (specific to the auditory paradigm) is to determine if sPCIst and sLZc discriminate between brain responses to auditory stimuli with eyes open vs. eyes closed. This aim is independent of aim 1, and although there is not a theoretically motivated prediction for this aim, it allows us to provide new evidence in the context of corresponding discrepancies reported in previous studies (Farnes et al., 2020; Mediano et al., 2024). To do so, we will look for differences in each complexity measure between eyes-open vs. eyes-closed conditions for one category each in two of the audio classes. Our two hypotheses for aim 2 are that there will be differences in sPCIst and sLZc between: 1) household-objects audio and household-objects audio eyes-closed; and 2) auditory noise and auditory noise eyes-closed.

1.4.3. Aim 3: visual vs. auditory modality

Our third aim is to determine if sPCIst and sLZc discriminate between brain responses to visual vs. auditory stimuli. This aim is also independent of aim 1. For this analysis, there are three levels of granularity: 1) between visual and auditory *paradigms*, such as between all visual stimuli and all auditory stimuli; 2) between visual and auditory *classes*, such as between natural images and natural audio; and 3) between categories *within* natural classes, such as between famous-people images and famous-names audio. Our five hypotheses for aim 3 are that there will be differences in sPCIst and sLZc between: 1) all images and all audio; 2) natural images and natural audio; 3) visual noise and auditory noise; 4) famous-people images and famous-names audio; and 5) household-objects images and household-objects audio.

1.4.4. Exploratory aim 1: dimensions of subjectivity

Our first exploratory aim (which complements aims 1-3) is to assess if any dimension(s) of subjective ratings (diversity/richness, unity/integratedness, meaningfulness, intelligibility/understandability, and intensity/vividness) correlate with any differences in the complexity measures. Even though many of the dimensions are conceptually similar to each other and/or some of the stimulus manipulations, we are making this aim exploratory, because for some of the dimensions in this context, they have either not been collected before (e.g., "intensity/vividness" and "unity/integratedness"), and/or because previous findings have not been sufficiently conclusive to inform specific predictions for them. Note also that this aim depends on finding differences in both the complexity measures and the subjective ratings, so it is not an independent aim.

1.4.5. Exploratory aim 2: reporting vs. no-reporting

Our second exploratory aim is to assess if sPCIst and sLZc reflect brain activity involved in reporting vs. not reporting. For this aim, we will aggregate trials across all stimulus modalities, classes, and categories into one reporting vs. no-reporting condition. The reporting condition is operationalized to include only the individual trials immediately following responses to either the subjective or behavioral task. Thus, if 'n' designates a trial in which a participant is tasked to respond, the reporting condition will include all 'n+1' trials. We are not treating the reporting trials themselves as the reporting condition, because on those trials, participants do not yet know that they will need to report by the time sufficient EEG data is collected post-stimulus (350-400 ms) to compute PCI and LZc. (This is by design in order to exclude any motor activity involved with reporting from the EEG window of interest). As such, we hypothesize that any residual reporting-related effect will be strongest on the trials immediately following the reporting trials, perhaps before it decays due to reports being tasked on only a random 33% subset of trials. Although we are implementing a behavioral task to control for attention and task demands, we are making this aim exploratory, because we are not aware of any previous findings or theories that justify predicting that sPCIst or sLZc reflect brain complexity associated with such task demands. This aim is however independent of aims 1-3.

1.4.6. Exploratory aim 3: correct vs. incorrect behavioral responses

Our third exploratory aim is to assess if sPCIst and sLZc reflect brain activity involved in correct vs. incorrect behavioral responses. Correct and incorrect responses will be operationalized based on whether or not participants choose the same category/class that was objectively presented. We will perform this analysis on the data aggregated by stimulus modality, class, and category. Although incorrect responses to the behavioral task could perhaps be interpreted as "unconscious" trials, we are making this aim exploratory, because this manipulation is not a strong control for this possibility, and we don't have an estimate for how many incorrect response trials to expect. This aim is also independent of aims 1-3.

1.4.7. Exploratory aim 4: spatial analysis

Our fourth exploratory aim is to assess if sPCIst and sLCz vary substantially across electrodes. In our primary analyses, we will compute these complexity measures on all electrodes, because we don't have sufficient theoretical nor empirical justification for focusing on a specific subset of electrodes. For this exploratory aim, we will compute these complexity measures for each electrode, inspect corresponding topographies, and compute the complexity measures on subsets that appear to give the best signal to noise ratio. Therefore, because it is not known whether these complexity measures systematically differ in certain subsets of electrodes, we are making this aim exploratory. This aim is also independent of aims 1-3.

1.4.8. Exploratory aim 5: time window analysis

Our fifth exploratory aim is to assess if sPCIst and sLCz vary substantially across time windows. In our primary analysis, we will limit the windows to the first 350 ms for auditory stimuli and 400 ms for visual stimuli based. This is based first on the fact that the default response duration for PCIst is 300 ms, but since PCI was developed based on TMS stimulation, we have added 50 ms and 100 ms to account for auditory and visual processing, respectively. Secondly, the relatively short duration is consistent with empirical findings that indicate that upon controlling for pre- and post-perceptual (cognitive) processes, the time course of sensory experiences themselves appear localizable in time to durations as early as ~120-200 ms (Dembski et al., 2021). For this exploratory aim, we will compute these complexity measures across various time windows, such as [0 ms, 200 ms], [200 ms, 400 ms], and [400 ms, 700 ms]. However, because it is not known whether these complexity measures systematically differ across these time windows, we are making this aim exploratory. This aim is also independent of aims 1-3.

1.4.9. Exploratory aim 6: EEG reference analysis

Our sixth exploratory aim is to assess if sPCIst and sLCz vary substantially at different latencies average- and mastoids-reference preprocessing methods. In our primary analysis, both the visual and auditory data will be re-referenced using the average of all of the electrodes (to be consistent), because mastoids-referencing has the potential to subtract out potentially relevant auditory processing activity. For this exploratory aim, we will re-reference the data to the average of the mastoids electrodes. However, because it is not known whether the reference method affects these complexity measures, we are making this aim exploratory. This aim is also independent of aims 1-3.

1.4.10. Exploratory aim 7: categorical relationships between classes

Our seventh exploratory aim is to assess if any relationships between stimulus classes (natural, blurred/vocoded, and noise) may be better characterized as categorical rather than linear. Although some studies support a linear structure across stimulus classes (Boly et al., 2015), other studies have indicated variability in the ordering across conditions, complexity measures, and sensory modalities (Bola et al., 2018; Mensen et al., 2017, 2018; Orłowski & Bola., 2023). Therefore, while we will use linear coding during data collection, based on both prior studies and our proof-of-concept analyses as a pragmatic modeling choice to improve model efficiency, we will also explore categorical coding to assess whether non-linear relationships exist in the data.

A summary of our nine research questions, 17 hypotheses, and supporting details is given in tables 1-3. All hypotheses are being made at the group level. ~~We are not presenting directional hypotheses for any of our three primary aims precisely because in previous work directions of~~

~~these effects have been inconsistent.~~ A summary of our seven ~~three~~ exploratory research questions is given in table 4.

Table 1. Summary of the three research questions and five hypotheses for the visual experimental paradigm.

Visual Paradigm						
Research Question	**Hypothesis**	**Sampling Plan**	**Analysis Plan**	**Hypothesis Test Sensitivity Rationale**	**Interpretation Given Different Outcomes**	**Theory That Could Be Shown Wrong**
1) Do sPCIst and sLZc discriminate between brain responses to **coarse-grained** differences in visual stimuli? (Aim 1)	There will be differences in sPCIst and sLZc (DVs) between: H1: Natural vs. blurred images (IV); H2: Natural images vs. visual noise (IV); H3: Blurred images vs. visual noise (IV); H0 (for H1-H4): There will not be differences in sPCIst and sLZc	Based on an attrition rate of 15%, we will aim to recruit a minimum of 30 and a maximum of 60 participants. The minimum target (25) is based on aiming to be the largest study of its kind to date. The maximum target (51) is based on the largest sample size for one of the measures from 10 power analyses conducted from three “pilot” datasets. Each power analysis (200 iterations) used the number of trials from the proposed study (540), an alpha level of 0.05, and 80% power. We will also compute Bayes factors	We will use Bayesian linear mixed effects models to determine if there are any differences in DVs between conditions. Each model will include fixed effects for condition/ interactions, and random/ varying effects for intercepts and slopes by participant to account for interindividual differences	We will assess whether 95% HPDIs for model coefficients cross 0, an approach commonly used to balance precision and confidence in Bayesian hypothesis testing. In addition, we will complement this method with Bayes factors and probability of direction analyses. We will interpret BF>3 as moderate evidence for the alternative hypothesis and BF<⅓ as moderate evidence for the null hypothesis. Consistent with our primary aims being non-directional, we will interpret probabilities of direction greater than	1) If we find evidence for differences in sPCIst and sLZc, this would suggest that these measures of complexity discriminate between brain responses to coarse/fine-grained visual stimuli (H1-H3); 2) If we do *not* find evidence for differences in sPCIst and sLZc, this would suggest (based on these data) that there is not strong support that these measures of complexity discriminate between brain responses to coarse/fine-grained visual stimuli, and/or that the effect was smaller than our effect size of interest	If outcome #2 occurs, theories that could be shown to be wrong would be any theory that supports that sPCIst and sLZc reflect differences in conscious *content*, such as integrated information theory (Albantakis et al., 2023), the entropic brain hypothesis (Carhart-Harris, 2018), and potentially the global neuronal workspace theory (Farisco & Changeux, 2023)

		(BF) after collecting data from each participant (after n=25) to determine if sufficient evidence for either the alternative hypothesis (BF>3) or null hypothesis (BF<⅓) is obtained prior to 51 subjects (for all 17 hypotheses)		97.5% as moderate evidence for an effect		
2) Do sPCIst and sLZc discriminate between brain responses to **fine-grained** differences in visual stimuli? (Aim 1)	There will be differences in sPCIst and sLZc between: H4: Famous-people vs. household-objects images	As above	As above	As above	As above	As above
3) Do sPCIst and sLZc discriminate between brain responses to **identical** visual stimuli? (Aim 1)	There will be differences in sPCIst and sLZc between: H5: Unrecognizable blurred images vs. recognizable blurred images	As above	As above	As above	As above	As above

Table 2. Summary of four research questions and seven hypotheses for the auditory experimental paradigm.

Auditory Paradigm						
Research Question	**Hypotheses**	**Sampling Plan**	**Analysis Plan**	**Hypothesis Test Sensitivity Rationale**	**Interpretation Given Alternative Outcome**	**Theory That Could Be Shown Wrong**
4) Do sPCIst	There will be	Based on an	We will use	We will	1) If we find	If outcome #2

and sLZc discriminate between brain responses to **coarse-grained** differences in auditory stimuli? (Aim 1)	differences in sPCIst and sLZc (DVs) between: H6: Natural vs. noise-vocoded audio (IV); H7: Natural audio vs. auditory noise (IV); H8: Noise-vocoded audio vs. auditory noise (IV); H0 (for H5-H12): There will not be differences in sPCIst and sLZc	attrition rate of 15%, we will aim to recruit a minimum of 30 and a maximum of 60 participants. The minimum target (25) is based on aiming to be the largest study of its kind to date. The maximum target (51) is based on the largest sample size for one of the measures from 10 power analyses conducted from three “pilot” datasets. Each power analysis (200 iterations) used the number of trials from the proposed study (540), an alpha level of 0.05, and 80% power. We will also compute Bayes factors (BF) after collecting data from each participant (after n=25) to determine if sufficient evidence for either the alternative hypothesis (BF>3) or null hypothesis (BF<⅓) is	Bayesian linear mixed effects models to determine if there are any differences in DVs between conditions. Each model will include fixed effects for condition/ interactions, and random/ varying effects for intercepts and slopes by participant to account for interindividual differences	assess whether 95% HPDIs for model coefficients cross 0, an approach commonly used to balance precision and confidence in Bayesian hypothesis testing. In addition, we will complement this method with Bayes factors and probability of direction analyses. We will interpret BF>3 as moderate evidence for the alternative hypothesis and BF<⅓ as moderate evidence for the null hypothesis. Consistent with our primary aims being non-directional, we will interpret probabilities of direction greater than 97.5% as moderate evidence for an effect	evidence for differences in sPCIst and sLZc, this would suggest that these measures of complexity discriminate between brain responses to coarse/fine-grained (including eyes open vs eyes closed) auditory stimuli (H5-H10); 2) If we do *not* find evidence for differences in sPCIst and sLZc, this would suggest (based on these data) that there is not strong support that these measures of complexity discriminate between brain responses to coarse/fine-grained (including eyes open vs eyes closed) auditory stimuli, and/or that the effect was smaller than our effect size of interest	occurs, theories that could be shown to be wrong would be any theory that supports that sPCIst and sLZc reflect differences in conscious *content*, such as integrated information theory (Albantakis et al., 2023), the entropic brain hypothesis (Carhart-Harris, 2018), and potentially the global neuronal workspace theory (Farisco & Changeux, 2023)

		obtained prior to 51 subjects (for all 17 hypotheses)				
5) Do sPCIst and sLZc discriminate between brain responses to **fine-grained** differences in auditory stimuli? (Aim 1)	There will be differences in sPCIst and sLZc between: H9: Famous-names vs. household-objects audio	As above	As above	As above	As above	As above
6) Do sPCIst and sLZc discriminate between brain responses to **identical** auditory stimuli? (Aim 1)	There will be differences in sPCIst and sLZc between: H10: Unrecognizable noise-vocoded audio vs. recognizable noise-vocoded audio	As above	As above	As above	As above	As above
7) Do sPCIst and sLZc discriminate between brain responses to auditory stimuli with **eyes-open vs. eyes-closed**? (Aim 2)	There will be differences in sPCIst and sLZc between: H11: Household-objects audio vs. household-objects audio eyes-closed; H12: Auditory noise vs. auditory noise eyes-closed	As above	As above	As above	As above	As above

Table 3. Summary of two research questions and five hypotheses for the visual vs. auditory analysis paradigm.

Visual vs. Auditory						
Research	**Hypotheses**	**Sampling**	**Analysis**	**Hypothesis**	**Interpretatio**	**Theory That**

Question		Plan	Plan	Test Sensitivity Rationale	n Given Alternative Outcome	Could Be Shown Wrong
8) Do sPCIst and sLZc discriminate between brain responses to **coarse-grained** visual vs. auditory stimuli? (Aims 1 and 3)	There will be differences in sPCIst and sLZc (DVs) between: H13: all images vs. all audio (IV); H14: Natural images vs. natural audio (IV); H15: Visual noise vs. auditory noise (IV); H0 (for H13-H17): There will not be differences in sPCIst and sLZc	Based on an attrition rate of 15%, we will aim to recruit a minimum of 30 and a maximum of 60 participants. The minimum target (25) is based on aiming to be the largest study of its kind to date. The maximum target (51) is based on the largest sample size for one of the measures from 10 power analyses conducted from three “pilot” datasets. Each power analysis (200 iterations) used the number of trials from the proposed study (540), an alpha level of 0.05, and 80% power. We will also compute Bayes factors (BF) after collecting data from each participant (after n=25) to determine if sufficient evidence for either the	We will use Bayesian linear mixed effects models to determine if there are any differences in DVs between conditions. Each model will include fixed effects for condition/ interactions, and random/ varying effects for intercepts and slopes by participant to account for interindividual differences	We will assess whether 95% HPDIs for model coefficients cross 0, an approach commonly used to balance precision and confidence in Bayesian hypothesis testing. In addition, we will complement this method with Bayes factors and probability of direction analyses. We will interpret BF>3 as moderate evidence for the alternative hypothesis and BF<⅓ as moderate evidence for the null hypothesis. Consistent with our primary aims being non-directional, we will interpret probabilities of direction greater than 97.5% as moderate evidence for an effect	1) If we find evidence for differences in sPCIst and sLZc, this would suggest that these measures of complexity discriminate between brain responses to coarse/fine-grained visual vs. auditory stimuli (H11-H15); 2) If we do *not* find evidence for differences in sPCIst and sLZc, this would suggest (based on these data) that there is not strong support that these measures of complexity discriminate between brain responses to coarse/fine-grained visual vs. auditory stimuli, and/or that the effect was smaller than our effect size of interest	If outcome #2 occurs, theories that could be shown to be wrong would be any theory that supports that sPCIst and sLZc reflect differences in conscious *content*, such as integrated information theory (Albantakis et al., 2023), the entropic brain hypothesis (Carhart-Harris, 2018), and potentially the global neuronal workspace theory (Farisco & Changeux, 2023)

		alternative hypothesis (BF>3) or null hypothesis (BF<⅓) is obtained prior to 51 subjects (for all 17 hypotheses)				
9) Do sPCIst and sLZc discriminate between brain responses to **fine-grained** visual vs. auditory stimuli? (Aims 1 and 3)	There will be differences in sPCIst and sLZc between: H16: Famous-people images vs. famous-names audio; H17: Household-objects images vs. household-objects audio	As above	As above	As above	As above	As above

Table 4. Summary of our seven ~~three~~ exploratory research questions.

Exploratory Questions
1) Do any **dimensions** of subjective experience correlate with differences in sPCIst and sLZc in response to coarse- and fine-grained visual and auditory stimuli?
2) Does **reporting vs. not reporting** correlate with differences in sPCIst and sLZc in response to coarse- and fine-grained visual and auditory stimuli?
3) Do **correct vs. incorrect** behavioral responses correlate with differences in sPCIst and sLZc in response to coarse- and fine-grained visual and auditory stimuli?
4) Do sPCIst and sLZc vary across **electrodes**?
5) Do sPCIst and sLZc vary across **time windows**?
6) Do sPCIst and sLZc vary across **EEG reference methods**?
7) Are any **relationships** between stimulus classes (natural, blurred/vocoded, and noise) **categorical**?

2. Methods

2.1. Sampling plan

To determine the sample size for the proposed study, we are using a combination of power analyses based on a proof-of-concept analysis (”pilot”) and a modified Sequential Bayes Factor Design (Schönbrodt & Wagenmakers, 2018). For the pilot study, we downloaded and analyzed three relevant EEG datasets from ERP Core (https://osf.io/thsqg/wiki/home), which is an open-source repository of datasets and preprocessing scripts for seven canonical ERP paradigms (Kappenman et al., 2021). For each of the three datasets, we computed sPCI and sLZc for each trial (in response to sensory stimuli) and looked for differences in each measure between corresponding conditions. To analyze the pilot data, we fit Bayesian linear mixed effects models to the sPCI and sLZc results. Detailed descriptions of the data, analyses, and results are provided in a supplementary file.

For the power analyses based on the pilot data, we used custom scripts (R version 4.2.0) to estimate all model coefficients across all three pilot datasets (10 power analyses in total). To conduct each power analysis, we simulated data (for a range of sample sizes) based on the coefficients and standard deviations for all fixed and random effects from the models we fit to the pilot data.[22] The simulated data (1200 iterations across the six final model coefficients for which there were effects) was based on the same number of trials of the proposed study (540). After fitting these models to the simulated data,[23] we computed the proportion of simulated p-values under an alpha level of 0.05 and based on a power of 80%. The final calculation of the sample size for the proposed study is based on the maximum sample size of five model coefficients for one of the measures (sPCI) across the three pilot datasets (51 subjects).

Although this sample size is somewhat underpowered for sLZc, we still consider it appropriate given its magnitude ~~(51 subjects)~~ and real-world constraints, because if this study yields significant results for only PCI, that would still be an economically significant finding in terms of a difference in sensitivity between PCI and LZc. However, we will ~~also~~ treat this sample size as an upper bound, because we will also compute Bayes factors for all 17 hypotheses after collecting data from each participant after n=25 to determine if sufficient evidence for either the alternative hypothesis (BF>3) or null hypothesis (BF<⅓) is obtained prior to 51 subjects. We will therefore stop data collection when BFs are conclusive for all 17 hypotheses, or upon reaching 51 subjects, whichever comes first. The minimum sample size of 25 is based on aiming to be the largest study of its kind to date. The target sample size is thus 25-51 subjects.

[22] Because we are proposing a paradigm that differs marginally from the paradigms used in the pilot data, we could not use the exact models of the proposed study in the power analyses since we did not have coefficients from the pilot results for all of the predictors and levels of each predictor of the proposed study. Nevertheless, the stimuli, tasks, and models of the pilot datasets are similar enough to the proposed study that we consider them a better approximation of the effect sizes of interest compared to previous studies, because no previous study has investigated PCI and LZc in a stimulus-evoked EEG context.

[23] To make the computational time of the power analyses tractable, we used non-Bayesian versions of the models.

Table 5. Sample-size results from the power analyses for sPCI and sLZc for all 10 model coefficients across all three pilot datasets (rounded up to the nearest number of participants). NA indicates no effect was found in the pilot dataset, thus no sample size was generated from the corresponding power analysis.

		sPCI	sLZc
Face-perception	Meaningfulness	51	79
	ObjectCategory	31	25
	Meaningfulness: ObjectCategory	29	NA
Visual-oddball	Condition	6	4
Auditory-oddball	Condition	48	NA

To mitigate attrition, we will schedule the visual and auditory sessions close in time. Based on previous studies conducted in our lab, we expect an attrition rate of ~15% (including due to noisy EEG and task performance), so we plan to recruit a minimum of 30 and a maximum of 60 participants. We will recruit these participants from the University and surrounding communities through research-participant recruiting platforms and word of mouth.

2.1.1. Participants

Participants will be 18-39 years of age, approximately equal-numbered male and female adults, and all have native English competency, normal color perception, normal or corrected-to-normal vision, normal hearing, and no history of neurological injury or disease. Participants will be compensated with research participation course credits or at a rate of $15/hour, whichever they prefer. Participants will be excluded based on having artifacts on more than 25% of trials, accuracy below 75%, or fewer than 50% of trials remaining in any individual condition (Kappenman et al., 2021).

2.2.2. Procedure

Two experimental sessions will be scheduled with each participant. At the start of each session, they will be given the choice to provide their informed written consent, after which if so, they will be given instructions about the study and a demographic questionnaire (including age, sex, language background, and education background). After the questionnaire, an EEG cap will be applied, and they will participate in a brief training session (details below) so that they can become comfortable with the style of the stimuli and tasks, and in particular the rating system for the subjective task. All research will be performed according to ethical standards as outlined in the Declaration of Helsinki.

2.2. Design

2.2.1. The visual stimuli

In the visual paradigm, we are using three classes of images inspired by Mensen et al. (2017): 1) natural ("natural images"); 2) blurred natural images ("blurred images"); and 3) images of randomly shuffled natural images ("visual noise"). The natural images are composed of two categories: 1) famous people ("famous-people images"); and 2) household objects ("household-objects images").[24] Each of the six categories/classes has 30 images for a total of 180 images. The design is 2x3, within-subjects. The stimuli are being sourced from https://www.listchallenges.com/200-most-famous-people-of-all-time/list/2 and https://www.wikipedia.org.

We are using image-blurring, because blurred images maintain some higher-order image statistics (e.g., shapes and colors) in the same range as the natural images but can be calibrated so as to not to be phenomenally distinguishable. Therefore, blurred images serve as controls for visual "complexity" while allowing for differences in phenomenology to manifest compared to natural stimuli. We will create blurred images for each of the natural images with custom scripts in Python (version 3.10.12) using Gaussian filters.

By contrast, and by design, randomly shuffling pixel values to produce something akin to visual noise preserves no higher-order image statistics (nor meaning). Therefore, it serves as a good baseline, because many consecutive images are essentially indistinguishable after gaps in time between consecutive presentations. We will create the visual noise with custom scripts in Python (version 3.10.12) by randomly shuffling the pixel values in each image. All visual stimuli can be accessed on our open science repository.

2.2.2. The visual tasks (subjective and behavioral)

Each trial will begin with black text that reads "Click + to proceed" displayed on a white background until the participant clicks anywhere on the screen with their mouse. After the click, the fixation cross will remain for 1 s to allow residual neural motor activity to dissipate, followed by the target image for 1 s. The visual stimuli will be presented at a size of approximately 10 degrees, using Presentation (Neurobehavioral Systems, Inc.), on a 24" monitor, with a 59 Hz refresh rate.

After a random 33% of the trials (per stimulus, per class/category), participants will provide five subjective ratings (via mouse input with their right hand) for their experience of the image according to the following five dimensions: 1) diversity/richness; 2) unity/integratedness; 3) meaningfulness; 4) intelligibility/understandability; and 5) intensity/vividness. A summary of all five dimensions, their operationalizations, examples, and rationales is given in Table 6, which will be explained to participants during a training period (excluding the rationales), detailed below. For the rating tasks, each of the dimensions will be presented consecutively and described onscreen with black text on a white background according to its operationalization. Each dimension will be rated by the participant according to the following 5-point scale: 1) very low, 2) low, 3) medium, 4) high, 5) very high. The order of the dimensions will be randomized and counterbalanced across trials.[25]

[24] We are using only two categories as opposed to seven from Mensen et al. (2017), because two categories should be sufficient to investigate within-class granularity. Furthermore, fewer categories allows us to repeat the stimuli and collect subjective ratings on a subset of the trials to control for attention and task demands while preventing the duration of the paradigm from increasing prohibitively.

[25] We aren't collecting ratings for familiarity, nor controlling for habituation/novelty and predictability, because Mensen et al. (2017) didn't find robust effects for those dimensions. Furthermore, this allows us to collect other ratings and add other manipulations, such as the eyes-open vs. eyes-closed condition.

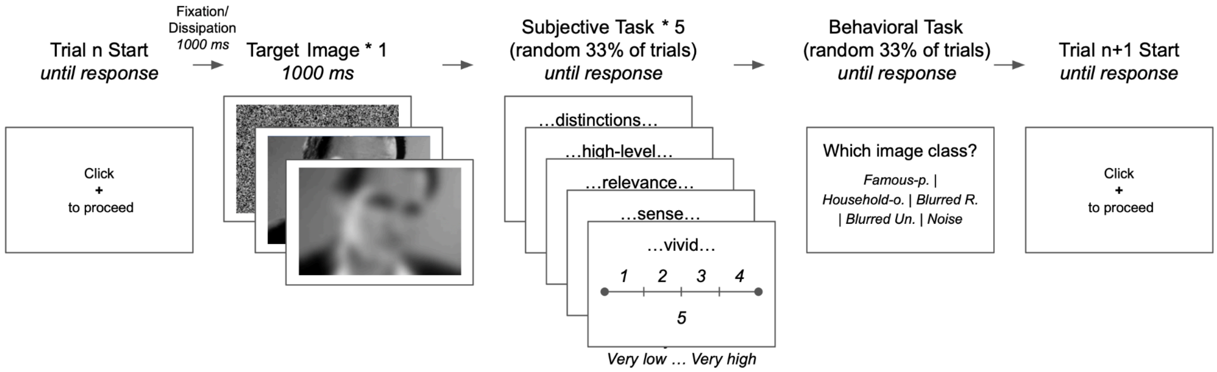

Figure 1. Overview of the visual tasks. Each trial will begin with black text that reads "Click + to proceed" displayed on a white background until the participant clicks the screen. After the click, the fixation cross will remain for 1 s to allow residual neural motor activity to dissipate, followed by the target image for 1 s (one of six categories/classes). After a random 33% of the trials (to control for task demands), participants will rate their experience from 1 to 5 according to five dimensions (consecutively): 1) diversity/richness; 2) unity/integratedness; 3) meaningfulness; 4) intelligibility/understandability; and 5) intensity/vividness. On the same 33% of the trials (to control for attention), participants will indicate which stimulus category/class they experienced (natural, blurred recognizable, blurred unrecognizable, or noisy). Each participant will complete 540 trials.

Table 6. Summary of the five subjective dimensions to be used by participants to rate their experiences of the stimuli.

Dimension	**Operationalization**	**Visual Examples**	**Auditory Examples**	**Rationale**
Diversity/richness	The number of recognizable *distinctions* in your experience of the image	An image with many shapes, objects, and other things that could generally be pointed to and described, as opposed to a blank image, or "nothingness"	A sound with many phonemes or other things that could generally be "pointed to" and described, as opposed to silence, or "nothingness"	A proposed phenomenological correlate of functional differentiation; similar dimensions were used in Mensen et al. (2017, 2018)
Unity/integratedness	How *high-level* your experience of the image is beyond just the low-level shapes and objects	An image of a rightside up face vs. an upside down face	A sound that comes together as a perceptual whole, such as a child saying "dada" to mean father, vs. having only (many) "disconnected" speech sounds/ phonemes, such as a child saying "tata"	A proposed phenomenological correlate of functional integration; not yet used by previous studies in the context of sensory-evoked conscious contents
Meaningfulness	Your sense of *relevance* of your experience of the	An image of a family member vs. a complete stranger,	A sound of a voice of a family member vs. a complete	A separate proposed phenomenological

	image	or an image of your first car vs. a car you do not recognize	stranger	correlate of the joint presence of functional differentiation and functional integration; roughly the dimension that many previous studies have purportedly been investigating (Bola et al., 2018; Boly et al., 2015; Mensen et al., 2017, 2018; Orlowski & Bola, 2023), which we will explore in aim 4
Intelligibility/ understandability	How much *sense* you can make of the image without having to try	An image of the Eiffel Tower vs. a jumbled mess of arbitrary shapes and lines that look like they could be rearranged to be the Eiffel Tower	A sound in a fluent language vs. a harsh whisper that is really hard to grasp	Expected to correlate very strongly with meaningfulness, but we are including it to explore cases in which it might not; used by previous studies such as Mensen et al. (2018)
Intensity/vividness	How "*sharp*", *vivid*, or *powerful* your experience of the image is	An image with a bright light or saturated color vs. a low light or muted color	A sound that is loud or high-pitched vs. quiet or low-pitched	A dimension of experience often reported to increase with psychedelics (Mediano et al., 2024; Schartner et al., 2017a), but which has not been investigated in the context of sensory-evoked conscious contents

On the same 33% of the trials for which participants provide subjective ratings (and for all image-blurred trials), participants will indicate (via mouse input with their right hand) which stimulus category/class they experienced: the natural category, the blurred class (and whether it is recognizable or unrecognizable), or the noisy class. Correct and incorrect responses will be operationalized based on whether or not they choose the same category/class that was objectively presented. We chose 33% because it should be sufficient to ensure that participants will be paying attention on most trials so that their brain activity reliably corresponds with their experiences. It will thus allow us to explore if there are any differences in brain complexity on trials immediately succeeding these responses (vs. all other trials), for example, due to participants' expectations of task demands/reporting (exploratory aim 2). It will also allow us to explore if there are differences in complexity on trials with incorrect responses (exploratory aim 3).

In order to collect at least one subjective rating for each image, we will show each stimulus three times for a total of 540 images, which will be divided into six blocks to give participants time to rest. The repetitions in the visual paradigm will allow for the blurred images to become recognizable on the second and third presentations. This allows us to completely control for low- and high-level stimulus complexity (which is identical between trials) while allowing for differences in phenomenology to manifest between recognizable and unrecognizable experiences.

The order of all images will be randomized across all blocks. On average, we estimate each rating/response to take approximately 2 seconds, and each block to take approximately 12.5 minutes (including a 1 minute break) for a total of 74 minutes. We anticipate the visual paradigm to take 1.75 hours in total, including time for informed consent, instructions, EEG setup, questionnaires, training, and EEG teardown. The visual and auditory paradigms will be randomized and counterbalanced across participants (e.g., with odd-numbered participants getting the visual paradigm first, and vice versa). The training session will include 1 instance of each stimulus class/category (not used in the actual task), for a total of 6 stimuli, presented 3 times each, for a total of 18 trials, 33% of which they will be tasked with the subjective and behavioral reports.

2.2.3. The auditory stimuli

In the auditory paradigm, we will translate each natural image into a word and record it being spoken by the same native English speaker in a sound attenuated booth at a sampling rate of 48 kHz ("natural audio"). The natural audio will be composed of two categories that mirror those of the visual paradigm: 1) natural images of famous people's names translated into speech ("famous-names audio"); 2) natural images of household objects translated into speech ("household-objects audio"). In addition to the natural audio, there will be two other putatively non-meaningful auditory classes (analogous to the visual paradigm): 1) noise-vocoded natural audio ("noise-vocoded audio"); and 2) and randomly shuffled natural audio ("auditory noise"). One of the natural audio categories (household objects) and one of the non-meaningful audio classes (white noise) will also be used for the eyes-closed condition ("household-objects audio eyes-closed" and "auditory noise eyes-closed"). Each of the eight categories/classes will have 30 audio clips for a total of 240 audio clips. The design is 2x3, within-subjects.

We are using noise-vocoding, because it preserves the overall amplitude envelope of the auditory signal (and thus some higher-order acoustic features) but removes the fine temporal structure (Scott et al., 2000; Shannon et al., 1995). As a result, noise-vocoded speech tends to be unrecognizable (like a harsh whisper) and thus not phenomenally distinguishable without training. As is the case with image-blurring in the visual paradigm, noise-vocoded audio serves as a control for auditory complexity while allowing for differences in meaning to manifest compared to natural stimuli. To create the noise-vocoded stimuli, we will use custom scripts in Matlab (version R2024b) that take each natural audio clip, divide the frequency space into eight different bands, extract the amplitude envelope of each band, and use white noise to replace the energy in each band. Finally, we will add the eight white-noise amplitude-modulated segments back together to yield the complete noise-vocoded signal (Shahin et al., 2018). We will use eight bands to balance initial un-recognizability and subsequent recognizability.

Auditory white noise, by contrast, and by design, preserves no higher-order acoustic features (e.g., pitch or formant structures). Therefore, as is the case with the images of random noise in the visual paradigm, it serves as a good baseline, because any two consecutive clips of white noise are essentially indistinguishable, especially after gaps in time between consecutive presentations. We will create the clips of auditory noise using custom scripts in Python (version 3.10.12). All auditory stimuli can be accessed on our open science repository

2.2.4. The auditory tasks (subjective and behavioral)

Each eyes-open trial will begin with the same black text as the visual paradigm that reads “Click + to proceed” displayed on a white background until the participant clicks anywhere on the screen with their mouse. After the click, the fixation cross will remain for 1 s to allow residual neural motor activity to dissipate, followed by the target audio clip for the duration of the recording. Each eyes-closed trial will begin with black text that reads “Close your eyes, and click to proceed after your eyes are closed” displayed a white background, followed by the target audio clip for the duration of the recording. The audio clips will be presented over Yamaha HS8 studio monitor speakers at 75 dB.

After a random 33% of the trials (per stimulus, per class/category), participants will provide subjective ratings (via mouse input with their right hand) for their experience of the audio clip according to the same five dimensions as the visual paradigm: 1) diversity/richness; 2) unity/integratedness; 3) meaningfulness; 4) intelligibility/understandability; and 5) intensity/vividness. Examples of the subjective dimensions for the auditory paradigm are given in Table 6, which will be explained to participants during a training period (excluding the rationales), detailed below. For the rating tasks, each of the dimensions will be presented consecutively and described onscreen with black text on a white background according to its operationalization. Each dimension will be rated by the participant according to the following 5-point scale: 1) very low, 2) low, 3) medium, 4) high, 5) very high. The order of the dimensions will be randomized and counterbalanced across trials.

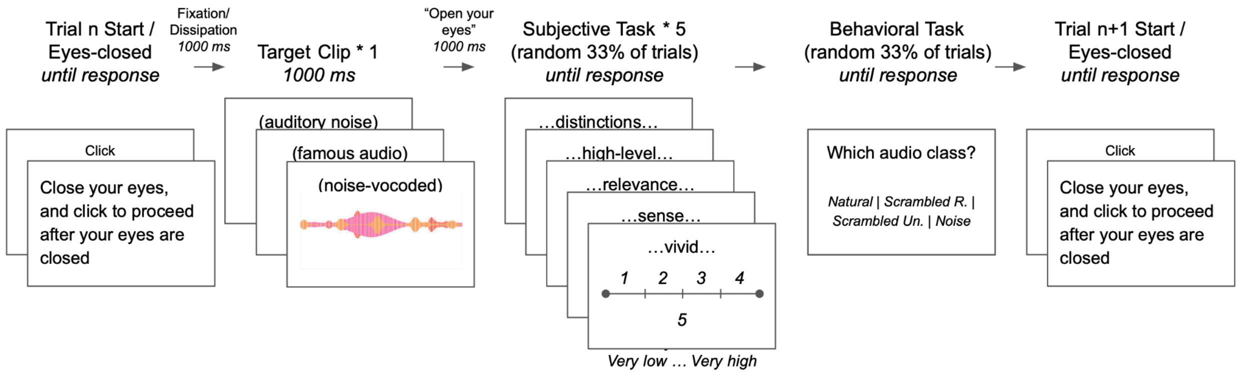

Figure 2. Overview of the auditory tasks. Each eyes-open trial will begin with onscreen text that reads "Click + to proceed". Each eyes-closed trial will begin with onscreen text that reads "Close your eyes, and click to proceed after your eyes are closed". After the participant clicks, the audio clip will be presented (1 of six categories/classes). After a random 33% of the trials (to control for task demands), participants will rate their experience from 1 to 5 according to five dimensions: 1) diversity/richness; 2) unity/integratedness; 3) meaningfulness; 4) intelligibility/understandability; and 5) intensity/vividness. On the same 33% of the trials (to control for attention), participants will indicate which stimulus category/class they experienced (natural, scrambled recognizable, scrambled unrecognizable, or noise). Each participant will complete 720 trials.

On the same 33% of the trials for which participants provide the subjective ratings (and for all noise-vocoded trials), participants will indicate (via mouse input with their right hand) which stimulus category/class they experienced: the natural category, the scrambled class (and whether or not it is recognizable or unrecognizable), or the noise class. Correct and incorrect responses will be operationalized based on whether or not they choose the same category/class that was objectively presented. This will again allow us to explore if there are any differences in brain complexity due to task demands/reporting, or on incorrect trials.

In order to collect at least one subjective rating for each stimulus, we will show each stimulus three times for a total of 720 audio clips, which will be divided into eight blocks to give participants time to rest (including separate blocks for the eyes-closed conditions). The repetitions in the auditory paradigm will allow for the noise-vocoded clips to become recognizable on the second and third presentations. This allows us to completely control for low- and high-level stimulus complexity (which is identical between trials) while allowing for differences in phenomenology to manifest between recognizable and unrecognizable experiences. In the eyes-closed blocks, the target auditory clips will be followed by the auditory instructions "Open your eyes" (whether or not those trials include the tasks).

The order of all audio clips will be randomized across all blocks within eyes-open and eyes-closed conditions. On average, we estimate each rating/response to take approximately 2 seconds, and each block to take approximately 12.5 minutes (including a 1 minute break) for a total of 101 minutes. We anticipate the auditory paradigm to take 2.25 hours in total, including time for informed consent, instructions, EEG setup, questionnaires, training, and EEG teardown. The visual and auditory paradigms will be randomized and counterbalanced across participants. The training session will include 1 instance of each stimulus class/category (not used in the actual task), for a total of 8 stimuli, presented 3 times each, for a total of 24 trials, 33% of which they will be tasked with the subjective and behavioral reports.

2.3. EEG acquisition

EEG signals will be recorded using a Biosemi ActiveTwo system (https://biosemi.com), with 64 electrodes arranged in an elastic cap according to the International 10/20 System. Two additional electrodes will be taped to the mastoids, and four additional electrodes will be taped above and below the eyes for electrooculogram recording (for a total of 70 electrodes). The EEG data will be sampled at a rate of 1024 Hz, and an online antialiasing low-pass filter will be applied during digitization. Electrode offsets (relative to the Common Mode Sense electrode) will be set to within ±20 μV for all channels (or as closely as possible). The recording room will also be sound-attenuated. Referencing will be performed offline.

2.4. EEG preprocessing

Preprocessing will be performed in Matlab (version R2024b) using open-source software plugins and custom scripts. Raw BDF files will be imported into EEGLAB (version 2024.2) (Delorme & Makeig, 2004) using the BioSig plugin (https://sourceforge.net/projects/biosig/), and the data from the blocks will be merged. An oscilloscope will be used to measure any delays in trigger timing, and this will be accounted for in the analysis if necessary. The data will be high-pass filtered (non-causal Butterworth impulse response function, half-amplitude cutoff of 0.1 Hz, 12 dB/oct roll-off) and low-pass filtered (non-causal Butterworth impulse response function, half-amplitude cutoff of 45 Hz, 12 dB/oct roll-off). The data will be downsampled to 256 Hz to make processing faster.

Artifact correction will be performed using EEGLAB's independent component analysis (ICA), using the Infomax algorithm (Bell & Sejnowski, 1995). Components associated with eyeblinks and horizontal eye movements will be removed by visual inspection. Channels with excessive levels of noise determined by visual inspection will be interpolated using EEGLAB's spherical interpolation algorithm. To be consistent, the visual and auditory data will be re-referenced using the average of all of the electrodes (because mastoids-referencing has the potential to subtract out potentially relevant auditory processing activity), and then segmented from -400 ms to 400 ms and baseline-corrected (using the average activity from -400 ms to 0 ms for artifact identification only, since windowing will also be done in Python before computing PCI and LZc). Segments of data that contain voltage-threshold artifacts (+-100 microvolts) will be flagged for subsequent removal.

2.5. PCI and LZc preprocessing

Before computing sPCIst and sLZc, we will exclude the trials that contain EEG artifacts. To compute sPCIst and sLZc, we will use the EEG recordings from all channels excluding the mastoid~~s~~ and electrooculogram channels.

2.6. Perturbational complexity index

Since PCIlz requires source-localization to generate the spatiotemporal matrix of significant activations, we will compute PCIst, which mitigates the need for source localization by utilizing principal components (and is faster to compute). We will compute sPCIst using Python (version 3.10.12) for every trial using freely available code (https://github.com/renzocom/PCIst) (Comolatti et al., 2019).

We will use the default PCIst parameter values for k=1.2, min_snr=1.1, max_var=99, embed=false, and n_steps=100. For the baseline window, we will use the default interval of approximately [-400 ms, -50 ms] (depending on exactly how the timepoints align based on the sample rate). For the response window, since the default values are based on a direct cortical perturbation via TMS, we will add 100 ms to the upper bound in the visual paradigm and 50 ms to the upper bound in the auditory paradigm to account for the additional time needed to process sensory stimuli. The resulting response windows will be approximately [0 ms, 400 ms] and [0 ms, 350 ms], respectively. After computing PCIst, we will exclude trials where PCIst=0.[26]

2.7. Lempel-Ziv complexity

Although most studies have shown that qualitative results for LZc remain consistent across various preprocessing techniques (Mediano et al., 2024; Schartner et al., 2015, 2017a), some studies have found minor discrepancies (Bola et al., 2018; Orlowski & Bola, 2023). Thus, we will compute three versions of LZc (all on the single-trial data) that have been used in related studies: 1) Lempel-Ziv complexity computed on the binarized data matrix channel by channel and then averaged (LZa) (Bola et al., 2018; Orlowski & Bola, 2023; Schartner et al., 2015); 2) Lempel-Ziv complexity computed on the binarized data matrix concatenated channel by channel (in "space") (LZcs) (Farnes et al., 2020; Schartner et al., 2015); and 3) Lempel-Ziv complexity computed on the binarized data matrix concatenated time-point by time-point (LZct) (Bola et al., 2018; Orlowski & Bola, 2023; Ort et al., 2023; Schartner et al., 2015).

To binarize the continuous EEG signal, we will use the mean of the absolute value (instantaneous amplitude) of the analytic (Hilbert-transformed) signal (Bola et al., 2018; Farnes et al., 2020; Orlowski & Bola, 2023; Ort et al., 2023; Schartner et al., 2015, 2017a), using an open-source Python (version 3.10.12) library (https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.hilbert.html). To most closely match the way that PCI analyzes the continuous EEG signal based on differences between response and baseline periods, for LZc – we will use the instantaneous amplitude of the baseline period of each channel/trial as the basis for binarizing the response period. To be consistent with how we are computing PCI, we will also compute LZc using the same baseline

[26] Based on our pilot data, we expect a small number of trials where PCIst=0, which we will exclude (and report). While the PCIst algorithm provides a signal-to-noise-ratio parameter, which when increased, yields more PCIst=0 values, when we performed a parameter sweep to try to minimize the number of PCIst=0 values, we were never able to completely eliminate them, so we left the parameter at its default value. We assume that PCIst=0 values indicate trials in which the PCIst algorithm is simply not sensitive enough to pick out the signal from the noise (as opposed to the implication that participants are losing consciousness), which is likely due to the subtle nature of sensory perturbations compared to TMS, the latter of which the PCIst algorithm was developed for.

and response windows for each trial. We will compute ~~Lempel-Ziv complexity~~ LZc using an open-source Python (version 3.10.12) library (https://github.com/Naereen/Lempel-Ziv_Complexity).

2.8. Analysis plan

Statistical analysis will be performed in R (version 4.4.2) using open-source software plugins and custom scripts. To compare results across participants and measures on a consistent scale, we will standardize (z-score) sPCIst and sLZc by subject. To determine whether there are any differences in sPCIst or sLZc between any of the conditions, we will use Bayesian linear mixed effects models (https://www.rdocumentation.org/packages/brms/versions/2.20.4).[27]

2.8.1. Primary aims

For the model specifications for our primary aims, we will include fixed effects for the Modality (auditory and visual), Class (natural, blurred/vocoded, and noise) and Category (household objects and famous-people) conditions, as well as for the interaction between the three predictors (to determine if the effects of Modality, Class, or Category vary between modalities, classes, or categories).[28] For Modality, we will code auditory as -0.5 and visual as 0.5 within a sum contrast coding scheme. For Class, since it is a 3-level factor, we will code noise as -1, blurred/vocoded as 0, and natural as 1 within a sum-contrast coding scheme, assuming an approximately linear relationship between the 3 levels, based on previous studies (Boly et al., 2015).[29] For Category, we will code household objects as -0.5 and famous-people as 0.5 within a sum contrast coding scheme.[30]

For most models, we will include random/varying effects for the intercept and slopes by participant and by item to account for differences between individual subjects and stimuli. We will specify weakly regularizing priors for all models,[31] and a Gaussian or skewed-normal likelihood (depending on which best fits the shape ~~distribution~~ of the data).[32] The following priors will be selected using domain knowledge and prior predictive checks: normal(0, 1) for the

[27] We will use Bayesian models because they provide a way to incorporate prior domain knowledge, and they permit more flexible comparisons to be made with the fitted model since posterior distributions are generated for all parameters of interest.

[28] If 3-way interactions or multiple random effects are too detrimental to model fit/performance, we can make the model simpler by removing them until model fit/performance improves.

[29] For our exploratory aim 7, which explores categorical coding to assess whether non-linear relationships exist in the data, we will use treatment coding with noise as the reference level to estimate direct pairwise contrasts between conditions. Additional comparisons (e.g., natural vs. blurred) will be assessed using post hoc tests (e.g., via emmeans).

[30] These coding schemes allow~~s~~ 1) the Modality coefficient to be interpreted as the unit change from the auditory to visual modalities; 2) the Class coefficient to be interpreted as the difference between the noise and scrambled stimuli, and as the difference between the scrambled and natural stimuli; and 3) the Category coefficient to be interpreted as the unit change from the household objects to famous people stimuli. The intercept will thus be interpreted as the grand mean of the auditory and visual modalities; of the noise, scrambled, and natural stimuli; and of the household-object and famous-people stimuli.

[31] We will specify weakly regularizing priors since we should have sufficient data to “overwhelm” the priors and don’t want to impose too many assumptions on the models.

[32] In our proof-of-concept analyses, we used normal likelihoods due to limited prior knowledge about the shape of the data. However, some measures and datasets exhibited skewed distributions. Therefore, in this study, we will base the choice on the best posterior model fit.

intercept; normal(0, 0.5) for the betas; normal(0, 1) for sigma; normal(0, 0.1) for sd; and lkj(2) for the correlation between the random intercepts and slopes.

The model specifications for our primary aims 1 and 3 (hypotheses 1-4, 6-9, and 13-17) will be:

*[sPCIst, sLZct, sLZcs, sLZa] ~ 1 + Modality * Class * Category +*
*(1 + Modality * Class * Category | subject_id) +*
*(1 + Modality * Class * Category | item_id)*

The interaction terms allow us to determine which level(s) of granularity any corresponding differences in complexity can be found, for example, for aim 3, which investigates if sPCIst and sLZc discriminate between brain responses to visual vs. auditory stimuli.

For the eyes-open vs. eyes-closed paradigm, we will remove the Modality predictor and add another predictor for the eyes-closed condition (EyesClosed). The model specifications for our primary aim 2 (hypotheses 11-12) will be:

*[sPCIst, sLZct, sLZcs, sLZa] ~ 1 + AudioClass * AudioCategory * EyesClosed +*
*(1 + AudioClass * AudioCategory * EyesClosed | subject_id) +*
*(1 + AudioClass * AudioCategory * EyesClosed | item_id)*

For the blurred images (hypothesis 5) and noise-vocoded audio (hypothesis 10), we will have a predictor for only whether or not the stimulus was reported as recognizable, and the models will be applied to only the blurred and vocoded data (individually). The model specifications will be:

[sPCIst, sLZct, sLZcs, sLZa] ~ 1 + Recognizable + (1 + Recognizable | subject_id) +
(1 + Recognizable | item_id)

As a complement to assessing whether 95% HPDIs for model coefficients cross 0, we will compute Bayes factors (to compare model evidence) and probabilities of direction (to evaluate the likelihood of positive or negative effects) using open-source libraries (https://cran.r-project.org/web/packages/bayestestR/index.html). We will interpret Bayes factors greater than 3 as moderate evidence for the alternative hypothesis and less than ⅓ as moderate evidence for the null hypothesis. Consistent with our primary aims being non-directional, we will interpret probabilities of direction greater than 97.5% as moderate evidence for an effect.

2.8.2. Exploratory aims

For ~~the three~~ exploratory ~~paradigms~~ aims 1-3, we will have predictors for each of the corresponding aims (dimensions of subjectivity, reporting vs. no-reporting, and correct vs. incorrect responses). For exploratory aim 1 (dimensions of subjectivity), the model specification will be:

[sPCIst, sLZct, sLZcs, sLZa] ~ 1 + Diversity + Unity + Meaningfulness + Intelligibility + Intensity +
(1 + Diversity + Unity + Meaningfulness + Intelligibility + Intensity | subject_id) +

(1 + Diversity + Unity + Meaningfulness + Intelligibility + Intensity | item_id)

In case of collinearity among the predictors, we will use principal component decomposition to reduce the dimensionality of the subjective dimensions, resulting in a smaller set of uncorrelated predictors that explain most of the variance in the original predictors.

To investigate whether any differences between eyes-open and eyes-closed conditions correspond with any differences in subjective ratings, we will include the EyesClosed predictor and its interactions with subjective dimensions in the following model.

*[sPCIst, sLZct, sLZcs, sLZa] ~ 1 + EyesClosed * (Diversity + Unity + Meaningfulness + Intelligibility + Intensity) +*
*(1 + EyesClosed * (Diversity + Unity + Meaningfulness + Intelligibility + Intensity) | subject_id) +*
*(1 + EyesClosed * (Diversity + Unity + Meaningfulness + Intelligibility + Intensity) | item_id)*

For exploratory aim 2 (reporting vs. no-reporting), the model specification will be:

[sPCIst, sLZct, sLZcs, sLZa] ~ 1 + Reporting + (1 + Reporting | subject_id) +
(1 + Reporting | item_id)

For exploratory aim 3 (correct vs. incorrect responses), the model specification will be:

[sPCIst, sLZct, sLZcs, sLZa] ~ 1 + Correct + (1 + Correct | subject_id) +
(1 + Correct | item_id)

Exploratory aims 4-6 may not involve statistical models, so we are not specifying models in advance. Exploratory aim 7 is captured by the model for primary aim 1 with the categorical coding modification

References

Albantakis, L., Barbosa, L., Findlay, G., Grasso, M., Haun, A. M., Marshall, W., ... & Tononi, G. (2023). Integrated information theory (IIT) 4.0: formulating the properties of phenomenal existence in physical terms. *PLoS Computational Biology, 19*(10), e1011465.

Bai, Y., Liang, Z., Li, X., Voss, L. J., & Sleigh, J. W. (2015). Permutation Lempel–Ziv complexity measure of electroencephalogram in GABAergic anaesthetics. *Physiological measurement, 36*(12), 2483.

Bayne, T., Frohlich, J., Cusack, R., Moser, J., & Naci, L. (2023). Consciousness in the cradle: on the emergence of infant experience. *Trends in cognitive sciences*.

Bayne, T., Hohwy, J., & Owen, A. M. (2016). Are there levels of consciousness?. *Trends in cognitive sciences, 20*(6), 405-413.

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.

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of cognitive neuroscience, 8*(6), 551-565.

Bertschinger, N., & Natschläger, T. (2004). Real-time computation at the edge of chaos in recurrent neural networks. *Neural computation, 16*(7), 1413-1436.

Block, N. (1995). On a confusion about a function of consciousness. Behavioral and brain sciences, 18(2), 227-247.

Bola, M., Orłowski, P., Baranowska, K., Schartner, M., & Marchewka, A. (2018). Informativeness of auditory stimuli does not affect EEG signal diversity. *Frontiers in psychology, 9*, 395515.

Boly, M., Coleman, M. R., Davis, M. H., Hampshire, A., Bor, D., Moonen, G., ... & Owen, A. M. (2007). When thoughts become action: an fMRI paradigm to study volitional brain activity in non-communicative brain injured patients. *Neuroimage, 36*(3), 979-992.

Boly, M., Garrido, M. I., Gosseries, O., Bruno, M. A., Boveroux, P., Schnakers, C., ... & Friston, K. (2011). Preserved feedforward but impaired top-down processes in the vegetative state. *Science, 332*(6031), 858-862.

Boly, M., Sasai, S., Gosseries, O., Oizumi, M., Casali, A., Massimini, M., & Tononi, G. (2015). Stimulus set meaningfulness and neurophysiological differentiation: a functional magnetic resonance imaging study. *PloS one, 10*(5), e0125337.

~~Boly, M., Seth, A. K., Wilke, M., Ingmundson, P., Baars, B., Laureys, S., ... & Tsuchiya, N. (2013). Consciousness in humans and non-human animals: recent advances and future directions. *Frontiers in psychology, 4*, 625.~~

Bötzel, K., & Grüsser, O. J. (1989). Electric brain potentials evoked by pictures of faces and non-faces: a search for “face-specific” EEG-potentials. *Experimental Brain Research, 77*, 349-360.

Canales-Johnson, A., Billig, A. J., Olivares, F., Gonzalez, A., Garcia, M. D. C., Silva, W., ... & Bekinschtein, T. A. (2020). Dissociable neural information dynamics of perceptual integration and differentiation during bistable perception. *Cerebral Cortex, 30*(8), 4563-4580.

Carhart-Harris, R. L. (2018). The entropic brain-revisited. *Neuropharmacology, 142*, 167-178.

Carhart-Harris, R. L., Leech, R., Hellyer, P. J., Shanahan, M., Feilding, A., Tagliazucchi, E., ...

& Nutt, D. (2014). The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs. *Frontiers in human neuroscience, 8*, 20.
Casali, A. G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K. R., ... & Massimini, M. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Science translational medicine, 5*(198), 198ra105-198ra105.
Cavanna, F., Muller, S., de la Fuente, L. A., Zamberlan, F., Palmucci, M., Janeckova, L., ... & Tagliazucchi, E. (2022). Microdosing with psilocybin mushrooms: a double-blind placebo-controlled study. *Translational psychiatry, 12*(1), 307.
Cheesman, J., & Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology/Revue canadienne de psychologie, 40*(4), 343.
Cocchi, L., Gollo, L. L., Zalesky, A., & Breakspear, M. (2017). Criticality in the brain: A synthesis of neurobiology, models and cognition. *Progress in neurobiology, 158*, 132-152.
Comolatti, R., Pigorini, A., Casarotto, S., Fecchio, M., Faria, G., Sarasso, S., ... & Casali, A. G. (2019). A fast and general method to empirically estimate the complexity of brain responses to transcranial and intracranial stimulations. *Brain stimulation, 12*(5), 1280-1289.
Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods, 134*(1), 9-21.
Dembski, C., Koch, C., & Pitts, M. (2021). Perceptual awareness negativity: a physiological correlate of sensory consciousness. *Trends in Cognitive Sciences*, 25(8), 660-670.
Dienes, Z., Altmann, G., Kwan, L., & Goode, A. (1995). Unconscious knowledge of artificial grammars is applied strategically. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(5), 1322.
Dulany, D. E. (1997). Consciousness in the explicit (deliberative) and implicit (evocative). In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 179–212). Lawrence Erlbaum Associates, Inc.
Engel, A. K., & Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends in cognitive sciences, 5*(1), 16-25.
Eriksen, C. W. (1960). Discrimination and learning without awareness: a methodological survey and evaluation. *Psychological review, 67*(5), 279.
Farisco, M., & Changeux, J. P. (2023). About the compatibility between the perturbational complexity index and the global neuronal workspace theory of consciousness. *Neuroscience of Consciousness, 2023*(1), niad016.
Farnes, N., Juel, B. E., Nilsen, A. S., Romundstad, L. G., & Storm, J. F. (2020). Increased signal diversity/complexity of spontaneous EEG, but not evoked EEG responses, in ketamine-induced psychedelic state in humans. *Plos one, 15*(11), e0242056.
Faugeras, F., Rohaut, B., Weiss, N., Bekinschtein, T. A., Galanaud, D., Puybasset, L., ... & Naccache, L. (2011). Probing consciousness with event-related potentials in the vegetative state. *Neurology, 77*(3), 264-268.
Fischer, C., Luaute, J., & Morlet, D. (2010). Event-related potentials (MMN and novelty P3) in

permanent vegetative or minimally conscious states. *Clinical neurophysiology, 121*(7), 1032-1042.
Fox, R., Todd, S., & Bettinger, L. A. (1975). Optokinetic nystagmus as an objective indicator of binocular rivalry. *Vision research, 15*(7), 849-853.
Frässle, S., Sommer, J., Jansen, A., Naber, M., & Einhäuser, W. (2014). Binocular rivalry: frontal activity relates to introspection and action but not to perception. *Journal of Neuroscience, 34*(5), 1738-1747.
Frohlich, J., Bayne, T., Crone, J. S., DallaVecchia, A., Kirkeby-Hinrup, A., Mediano, P. A., ... & Preissl, H. (2023). Not with a "zap" but with a "beep": measuring the origins of perinatal experience: Origins of perinatal experience. *NeuroImage, 120057*.
Frohlich, J., Toker, D., & Monti, M. M. (2021). Consciousness among delta waves: a paradox?. *Brain, 144*(8), 2257-2277.
Gosseries, O., Di, H., Laureys, S., & Boly, M. (2014). Measuring consciousness in severely damaged brains. *Annual Review of Neuroscience, 37*, 457-478.
Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1, pp. 1969-2012). New York: Wiley.
~~Hohwy, J. (2009). The neural correlates of consciousness: new experimental approaches needed?. *Consciousness and cognition, 18*(2), 428-438.~~
Holland, P. W. (1986). Statistics and causal inference. *Journal of the American statistical Association*, *81*(396), 945-960.
Höller, Y., Bergmann, J., Kronbichler, M., Crone, J. S., Schmid, E. V., Golaszewski, S., & Ladurner, G. (2011). Preserved oscillatory response but lack of mismatch negativity in patients with disorders of consciousness. *Clinical Neurophysiology, 122*(9), 1744-1754.
Honey, C., Kirchner, H., & VanRullen, R. (2008). Faces in the cloud: Fourier power spectrum biases ultrarapid face detection. *Journal of vision, 8*(12), 9-9.
Johnson, R. W., & Shore, J. E. (1984). Relative-entropy minimization with uncertain constraints—theory and application to Spectrum analysis. *NRL Memorandum Report, 5478*.
Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*(11), 4302-4311.
Kappenman, E. S., Farrens, J. L., Zhang, W., Stewart, A. X., & Luck, S. J. (2021). ERP CORE: An open resource for human event-related potential research. *NeuroImage, 225*, 117465.
Kissin, I. (2000). Depth of anesthesia and bispectral index monitoring. Anesthesia & Analgesia, 90(5), 1114-1117.
Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness: progress and problems. *Nature Reviews Neuroscience, 17*(5), 307-321.
Koculak, M., & Wierzchoń, M. (2022). How much consciousness is there in complexity?. *Frontiers in Psychology, 13*, 983315.
Kotchoubey, B. (2005). Event-related potential measures of consciousness: two equations with three unknowns. *Progress in brain research, 150*, 427-444.
~~Laureys, S. (2005). The neural correlate of (un) awareness: lessons from the vegetative state. *Trends in cognitive sciences, 9*(12), 556-559.~~

Lempel, A., & Ziv, J. (1976). On the complexity of finite sequences. I*EEE Transactions on information theory, 22*(1), 75-81.
Leopold, D. A., Fitzgibbons, J. C., & Logothetis, N. K. (1995). The role of attention in binocular rivalry as revealed through optokinetic nystagmus.
Logothetis, N. K., & Schall, J. D. (1990). Binocular motion rivalry in macaque monkeys: eye dominance and tracking eye movements. *Vision research, 30*(10), 1409-1419.
Makeig, S., Bell, A., Jung, T. P., & Sejnowski, T. J. (1995). Independent component analysis of electroencephalographic data. *Advances in neural information processing systems, 8*.
Massimini, M., Boly, M., Casali, A., Rosanova, M., & Tononi, G. (2009). A perturbational approach for evaluating the brain's capacity for consciousness. *Progress in brain research, 177*, 201-214.
Mediano, P. A., Rosas, F. E., Timmermann, C., Roseman, L., Nutt, D. J., Feilding, A., ... & Carhart-Harris, R. L. (2024). Effects of external stimulation on psychedelic state neurodynamics. *ACS Chemical Neuroscience, 15*(3), 462-471.
Mensen, A., Marshall, W., Sasai, S., & Tononi, G. (2018). Differentiation analysis of continuous electroencephalographic activity triggered by video clip contents. *Journal of Cognitive Neuroscience, 30*(8), 1108-1118.
Mensen, A., Marshall, W., & Tononi, G. (2017). EEG differentiation analysis and stimulus set meaningfulness. *Frontiers in psychology, 8*, 274988.
Monti, M. M., Vanhaudenhuyse, A., Coleman, M. R., Boly, M., Pickard, J. D., Tshibanda, L., ... & Laureys, S. (2010). Willful modulation of brain activity in disorders of consciousness. *New England journal of medicine, 362*(7), 579-589.
Murray, C. H., Frohlich, J., Haggarty, C. J., Tare, I., Lee, R., & de Wit, H. (2024). Neural complexity is increased after low doses of LSD, but not moderate to high doses of oral THC or methamphetamine. *Neuropsychopharmacology, 1-9*.
Nagel, Thomas. “What Is It Like to Be a Bat?” *The Philosophical Review, vol. 83*, no. 4, 1974, pp. 435–50.
Ní Mhuircheartaigh, R., Warnaby, C., Rogers, R., Jbabdi, S., & Tracey, I. (2013). Slow-wave activity saturation and thalamocortical isolation during propofol anesthesia in humans. *Science translational medicine, 5*(208), 208ra148-208ra148.
Oizumi, M., Ishii, T., Ishibashi, K., Hosoya, T., & Okada, M. (2010). Mismatched decoding in the brain. *Journal of Neuroscience, 30*(13), 4815-4826.
Oizumi, M., Yanagawa, T., Amari, S. I., Tsuchiya, N., & Fujii, N. (2012). Measuring the level of consciousness based on the amount of integrated information. F*rontiers in Human Neuroscience, 75*.
Orłowski, P., & Bola, M. (2023). Sensory modality defines the relation between EEG Lempel–Ziv diversity and meaningfulness of a stimulus. *Scientific Reports, 13*(1), 3453.
Ort, A., Smallridge, J. W., Sarasso, S., Casarotto, S., von Rotz, R., Casanova, A., ... & Vollenweider, F. X. (2023). TMS-EEG and resting-state EEG applied to altered states of consciousness: oscillations, complexity, and phenomenology. *Iscience, 26*(5).
Overgaard, M. (2015). The challenge of measuring consciousness. *Behavioural methods in consciousness research, 7-19*.
~~Overgaard, M., & Overgaard, R. (2010). Neural correlates of contents and levels of consciousness. *Frontiers in psychology, 1*, 164.~~

Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S., & Pickard, J. D. (2006). Detecting awareness in the vegetative state. *Science, 313*(5792), 1402-1402.

Pal, D., Li, D., Dean, J. G., Brito, M. A., Liu, T., Fryzel, A. M., ... & Mashour, G. A. (2020). Level of consciousness is dissociable from electroencephalographic measures of cortical connectivity, slow oscillations, and complexity. *Journal of Neuroscience, 40*(3), 605-618.

Pearl, J., & Verma, T. S. (1995). A theory of inferred causation. In *Studies in Logic and the Foundations of Mathematics*(Vol. 134, pp. 789-811). Elsevier.

Pessoa, L., Japee, S., Sturman, D., & Ungerleider, L. G. (2006). Target visibility and visual awareness modulate amygdala responses to fearful faces. *Cerebral cortex, 16*(3), 366-375.

Pincus, S. M. (1991). Approximate entropy as a measure of system complexity. *Proceedings of the national academy of sciences, 88*(6), 2297-2301.

Pitts, M. A., Metzler, S., & Hillyard, S. A. (2014). Isolating neural correlates of conscious perception from neural correlates of reporting one's perception. *Frontiers in psychology, 5,* 105150.

Plourde, G., Stapells, D. R., & Picton, T. W. (1991). The human auditory steady-state evoked potentials. *Acta Oto-Laryngologica, 111*(sup491), 153-160.

Radhakrishnan, N., & Gangadhar, B. N. (1998). Estimating regularity in epileptic seizure time-series data. IEEE engineering in medicine and biology magazine, 17(3), 89-94.

Railo, H., Koivisto, M., & Revonsuo, A. (2011). Tracking the processes behind conscious perception: a review of event-related potential correlates of visual consciousness. *Consciousness and cognition, 20*(3), 972-983.

Rampersad, S. E., & Mulroy, M. F. (2005). A case of awareness despite an “adequate depth of anesthesia” as indicated by a Bispectral Index® monitor. Anesthesia & Analgesia, 100(5), 1363-1364.

Ramsøy, T. Z., & Overgaard, M. (2004). Introspection and subliminal perception. *Phenomenology and the cognitive sciences, 3*, 1-23.

Rosanova, M., Gosseries, O., Casarotto, S., Boly, M., Casali, A. G., Bruno, M. A., ... & Massimini, M. (2012). Recovery of cortical effective connectivity and recovery of consciousness in vegetative patients. *Brain, 135*(4), 1308-1320.

Sandberg, K., & Overgaard, M. (2015). Using the perceptual awareness scale (PAS). *Behavioral methods in consciousness research*, 181-196.

Sandberg, K., Timmermans, B., Overgaard, M., & Cleeremans, A. (2010). Measuring consciousness: is one measure better than the other?. *Consciousness and cognition, 19*(4), 1069-1078.

Sanders, R. D., Tononi, G., Laureys, S., Sleigh, J. W., & Warner, D. S. (2012). Unresponsiveness ≠ unconsciousness. *The Journal of the American Society of Anesthesiologists, 116*(4), 946-959.

Sarasso, S., Boly, M., Napolitani, M., Gosseries, O., Charland-Verville, V., Casarotto, S., ... & Massimini, M. (2015). Consciousness and complexity during unresponsiveness induced by propofol, xenon, and ketamine. *Current Biology, 25*(23), 3099-3105.

Sarasso, S., Casali, A. G., Casarotto, S., Rosanova, M., Sinigaglia, C., & Massimini, M. (2021). Consciousness and complexity: a consilience of evidence. *Neuroscience of consciousness, 2021*(2), niab023.

Schartner, M. M., Carhart-Harris, R. L., Barrett, A. B., Seth, A. K., & Muthukumaraswamy, S. D. (2017a). Increased spontaneous MEG signal diversity for psychoactive doses of ketamine, LSD and psilocybin. *Scientific reports, 7(*1), 46421.

Schartner, M. M., Pigorini, A., Gibbs, S. A., Arnulfo, G., Sarasso, S., Barnett, L., ... & Barrett, A. B. (2017b). Global and local complexity of intracranial EEG decreases during NREM sleep. *Neuroscience of consciousness, 2017*(1), 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), e0133532.

Schönbrodt, F. D., & Wagenmakers, E. J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic bulletin & review*, 25(1), 128-142.

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, *123*(12), 2400-2406.

Seth, A. K., & Bayne, T. (2022). Theories of consciousness. N*ature Reviews Neuroscience, 23*(7), 439-452.

Seth, A. K., Dienes, Z., Cleeremans, A., Overgaard, M., & Pessoa, L. (2008). Measuring consciousness: relating behavioural and neurophysiological approaches. *Trends in cognitive sciences, 12*(8), 314-321.

Shahin, A. J., Shen, S., & Kerlin, J. R. (2017). Tolerance for audiovisual asynchrony is enhanced by the spectrotemporal fidelity of the speaker's mouth movements and speech. *Language, cognition and neuroscience, 32*(9), 1102-1118.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, *270*(5234), 303-304.

Sigl, J. C., & Chamoun, N. G. (1994). An introduction to bispectral analysis for the electroencephalogram. *Journal of clinical monitoring, 10*, 392-404.

Sinitsyn, D. O., Poydasheva, A. G., Bakulin, I. S., Legostaeva, L. A., Iazeva, E. G., Sergeev, D. V., ... & Piradov, M. A. (2020). Detecting the potential for consciousness in unresponsive patients using the perturbational complexity index. *Brain Sciences, 10*(12), 917.

Sitt, J. D., King, J. R., El Karoui, I., Rohaut, B., Faugeras, F., Gramfort, A., ... & 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.

Smyth, A. C., & Shanks, D. R. (2008). Awareness in contextual cuing with extended and concurrent explicit tests. *Memory & Cognition, 36*, 403-415.

Tassi, P., & Muzet, A. (2001). Defining the states of consciousness. Neuroscience & *Biobehavioral Reviews, 25*(2), 175-191.

Timmermann, C., Roseman, L., Schartner, M., Milliere, R., Williams, L. T., Erritzoe, D., ... & Carhart-Harris, R. L. (2019). Neural correlates of the DMT experience assessed with multivariate EEG. *Scientific reports, 9*(1), 16324.

Toker, D., Pappas, I., Lendner, J. D., Frohlich, J., Mateos, D. M., Muthukumaraswamy, S., ... & D'Esposito, M. (2022). Consciousness is supported by near-critical slow cortical electrodynamics. *Proceedings of the National Academy of Sciences, 119*(7), e2024455119.

Tononi, G., & Edelman, G. M. (1998). Consciousness and complexity. *Science, 282*(5395), 1846-1851.

Tononi, G., Sporns, O., & Edelman, G. M. (1994). A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Sciences, 91*(11), 5033-5037.

Tsuchiya, N., Wilke, M., Frässle, S., & Lamme, V. A. (2015). No-report paradigms: extracting the true neural correlates of consciousness. *Trends in cognitive sciences, 19*(12), 757-770.

Tzovara, A., Simonin, A., Oddo, M., Rossetti, A. O., & De Lucia, M. (2015). Neural detection of complex sound sequences in the absence of consciousness. *Brain, 138*(5), 1160-1166.

Vanhaudenhuyse, A., Noirhomme, Q., Tshibanda, L. J. F., Bruno, M. A., Boveroux, P., Schnakers, C., ... & Boly, M. (2010). Default network connectivity reflects the level of consciousness in non-communicative brain-damaged patients. *Brain, 133*(1), 161-171.

Velmans, M. (2009). How to define consciousness. *Journal of Consciousness Studies, 16*(5), 139–156.

Wang, J., Noh, G. J., Choi, B. M., Ku, S. W., Joo, P., Jung, W. S., ... & Lee, H. (2017). Suppressed neural complexity during ketamine-and propofol-induced unconsciousness. *Neuroscience letters, 653*, 320-325.

Warnaby, C. E., Sleigh, J. W., Hight, D., Jbabdi, S., & Tracey, I. (2017). Investigation of slow-wave activity saturation during surgical anesthesia reveals a signature of neural inertia in humans. *Anesthesiology, 127*(4), 645-657.

Zhang, X. S., Roy, R. J., & Jensen, E. W. (2001). EEG complexity as a measure of depth of anesthesia for patients. *IEEE transactions on biomedical engineering, 48*(12), 1424-1433.

Ziv, J., & Lempel, A. (1977). A universal algorithm for sequential data compression. *IEEE Transactions on information theory, 23*(3), 337-343.