The proposed study aims to investigate measures of EEG signal diversity and complexity as potential correlates of the content of consciousness. I think this might be an important and timely study since currently we have extensive evidence that these measures constitute correlates of the conscious state (i.e. the capacity to have an experience), but mixed evidence whether they can also reflect the content of consciousness (or some dimensions of experience). Below I provide some comments and suggestions which, in my opinion, might further increase the impact of the study.

Introduction: The first part of the paper is not a classic introduction but rather an extensive review of previous literature. The authors have nicely identified gaps and inconsistencies in previous research. However, section 1.3.3 is rather speculative, so maybe move it to discussion? Especially that it ends “This study will not explicitly investigate this possibility, but it would be an interesting topic for future investigation”.

Hypotheses: when presenting and discussion hypotheses the authors do not indicate the direction of the effect, for instance “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) animal images and animal audio; and 5) household-objects images and household-objects audio”. Can you please clarify why you do not present directional hypotheses?

P. 16 “Our second exploratory aim is to assess if sPCIst and sLZc reflect brain activity involved in reporting vs. not reporting.” - what does “reporting” mean in this context? How will it be operationalized?

Stimuli: Physical (low-level) differences in the visual stimuli used - using images of animals and household objects will most likely result in differences in physical features (e.g. spatial frequency) between these categories (and this will probably be reflected also in phase-scrambled versions). Will such differences be controlled, or do you assume they are of no relevance?

Analysis:
“Finally, a low-pass filter will be applied using a non-causal Butterworth impulse response function with a half-amplitude cutoff of 45 Hz and 12 dB/oct roll-off.” - ideally, a low-pass filter should be employed before downsampling to avoid the aliasing effect.

“Before computing sPCIst and sLZc, we will exclude the trials that contain EEG artifacts.” - this is described in the previous section so maybe no need to state again (it suggests that some subsequent data cleaning will be performed before calculating these measures).

I think that more details regarding calculations of sLZc need to be provided. Will these measures be calculated on single-trial data? Which time-window will be used (0-400, similarly as for sPCIst?)?

“To binarize the continuous EEG signal, we will use the mean of the absolute value (instantaneous amplitude) of the analytic (Hilbert-transformed) signal” - mean over a given trial/channel? How will this mean be calculated?

“After computing PCIst, we will exclude trials where PCIst=0.” - what might be the cause of the measure = 0?

A more general comment regarding the analysis is that we do not know how PCI/LZ measures depend on preprocessing steps.
One such step is re-referencing - here the authors decide to re-reference to the average signal. Is there any rationale behind using an average rather than mastoids? I would also suggest doing a control analysis with the signal referenced to mastoids.
Second, at the moment the spatial dimension is completely disregarded in the analysis. From my experience electrodes closer to the midline (3, z, and 4 line) typically exhibit lower diversity, while electrodes closer to the jaw and neck muscles (7, 8 line) exhibit higher diversity, just because they record more artefacts (which artificially increase signal diversity). Thus, here the authors might consider two things - first, to inspect topographies of diversity measures (create plots for each condition, but also differential plots with one conduction subtracted from the other) and, second, to conduct the analysis on a subset of electrodes closer to the midline (and excluding the typically most noisy electrodes). Both might be investigated and tested in the pilot data.

Proof of concept: The authors present results of a “proof of concept” analysis, which I think should be evaluated very positively - in the final version these can probably be presented in the supplementary material rather than the main manuscript.

Minor comments
P. 13 “After each trial, participants will provide subjective ratings for their experience of the stimulus according to the following five dimensions…” - but this will be done in ⅓ of trials, right?

P. 14 - “. This aim allows us to provide evidence not provided by previous studies (Mensen et al., 2017).” - please specify what kind of evidence.

Michał Bola, PhD