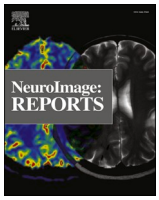
Neuroimage: Reports 2 (2022) 100089

Contents lists available at ScienceDirect

Neuroimage: Reports

journal homepage: www.sciencedirect.com/journal/neuroimage-reports



EEG microstates: Functional significance and short-term test-retest reliability

Elena Antonova [a,b,d,*], Martin Holding [b], Ho Chak Suen [b], Alex Sumich [c], Reinoud Maex [d], Chrystopher Nehaniv [e,d]

[a] Division of Psychology, Department of Life Sciences, College of Health, Medicine and Life Sciences & Centre for Cognitive Neuroscience, Brunel University London, Kingston Lane, Uxbridge, UB8 3PH, United Kingdom
[b] Department of Psychology, Institute of Psychiatry Psychology & Neuroscience, King's College London, De Crespigny Park, London, SE5 8AF, United Kingdom
[c] Department of Psychology, School of Social Sciences, Nottingham Trent University, 50 Shakespeare Street, Nottingham, NG1 4FQ, United Kingdom
[d] Royal Society Wolfson Biocomputation Research Laboratory, Centre for Computer Science and Informatics Research, University of Hertfordshire, College Lane, Hatfield, AL10 9AB, United Kingdom
[e] Departments of Systems Design Engineering and Electrical & Computer Engineering, University of Waterloo, 200 University Avenue W, Waterloo, Ontario, N2L 3G1, Canada

ARTICLE INFO

Keywords:
EEG microstates
Resting-state EEG
Resting-state networks
Neuroimaging
Biomarkers
Mind-wandering

ABSTRACT

EEG microstates, reflecting discrete topographical organization of the EEG signal power, may have clinical relevance; however, their functional significance and test-retest reliability remain unclear. To investigate the functional significance of the canonical EEG microstate classes and their pairwise transitions, and to establish their within-session test-retest reliability, we recorded 36-channel EEGs in 20 healthy volunteers during three eyes-closed conditions: mind-wandering, verbalization (silently repeating the word 'square' every 2 s), and visualization (visualizing a square every 2 s). Each condition lasted 3 min and the sequence of three conditions was repeated four times (two runs of two sequence repetitions). The participants' alertness and their sense of effort during the experiment were rated using visual-analogue scales. The EEG data were 2–20 Hz bandpass-filtered and analysed into the four canonical microstate classes: A, B, C, and D. EEG microstate classes C and D were persistently more dominant than classes A and B in all conditions. Of the first-order microstate parameters, average microstate duration was most reliable. The duration of class D microstate was longer during mind-wandering (106.8 ms) than verbalization (102.2 ms) or visualization (99.8 ms), with a concomitantly higher coverage (36.4% vs. 34.7% and 35.2%), but otherwise there was no clear association of the four microstate classes to particular mental states. The test-retest reliability was higher at the beginning of each run, together with higher average alpha power and subjective ratings of alertness. Only the transitions between classes C and D (from C to D in particular) were significantly higher than what would be expected from the respective microstates' occurrences. The transition probabilities, however, did not distinguish between conditions, and their test-retest reliability was overall lower than that of the first-order parameters such as duration and coverage. Further studies are needed to establish the functional significance of the canonical EEG microstate classes. This might be more fruitfully achieved by looking at their complex syntax beyond pairwise transitions. To ensure greater test-retest reliability of microstate parameters, study designs should allow for shorter experimental runs with regular breaks, particularly when using EEG microstates as clinical biomarkers.

1. Introduction

EEG microstates have become an established method for investigating the spatiotemporal dynamics of large-scale brain networks at rest and during cognitive information processing (review, Michel and Koenig, 2018). First described by Lehmann et al. (1987), EEG microstates are topographic maps of the momentary distributions in electric potential that remain quasi-stable in a millisecond time range (approx.

* Corresponding author. Division of Psychology, Department of Life Sciences, College of Health, Medicine and Life Sciences & Centre for Cognitive Neuroscience, Brunel University London, Kingston Lane, Uxbridge, UB8 3PH, United Kingdom.
E-mail address: elena.antonova@brunel.ac.uk (E. Antonova).

https://doi.org/10.1016/j.ynirp.2022.100089
Received 24 June 2021; Received in revised form 23 January 2022; Accepted 13 February 2022
Available online 12 March 2022

80–150 ms on average) and are separated by rapid spatial configuration changes. They can be obtained from event-related or resting-state data and further analysed into typological classes and their parameters, including: *duration* (the average length of time a microstate remains stable), *occurrence* (the average number of times per second the microstate is dominant during the total recording time), *coverage* (the fraction of total recording time that the microstate is dominant), and *syntax* (microstate sequence patterns). Understanding the functional significance and reliability of EEG microstates and their sequences is essential due to the emerging evidence of their utility as biomarkers for psychological and neurological disorders (review, Khanna et al., 2015).

Most resting-state EEG studies (e.g. Koenig et al., 1999; Kindler et al., 2011; Lehmann et al., 2005) reveal four topographic microstate classes, accounting for 70–80% of the variance. Two of these classes have left-right asymmetric topography: class A (right-frontal to left-posterior orientation) and class B (left-frontal to right-posterior orientation); and two have symmetric topography: class C (frontal to occipital orientation) and class D (frontal to occipital orientation, but with a more midline frontal maximum than class C) (See Fig. 1a). These four EEG microstate classes are referred to as *canonical* in the literature.

The functional significance of the four canonical microstates and their relationship with large-scale brain networks have become questions of growing research interest. The simultaneous EEG-fMRI study by Britz et al. (2010), investigating the relationship between the four canonical microstates and fMRI resting state networks, observed class A to correlate with *negative* BOLD signal in areas implicated in phonological processing including auditory and language comprehension areas in the left temporal lobe; class B with *negative* BOLD signal in areas implicated in visual processing and imagery including bilateral occipital cortex and cuneus; class C with *positive* BOLD signal in saliency (anterior cingulate) and interoceptive (insula) processing areas; and class D with *negative* BOLD signal in networks associated with right-lateralized dorsal attentional networks.

Pascual-Marqui et al. (2014) used exact Low Resolution Electromagnetic Tomography (eLORETA) to determine the cortical distributions of electric neuronal activity generating the microstates and found all four microstates to have common posterior cingulate generators, while microstate class A additionally included activity in the left occipital/parietal, class B in right occipital/parietal, and class C in the anterior cingulate cortices. Effective connectivity modelling performed in the same study identified the posterior cingulate cortex as an important hub in microstate generation, sending alpha and beta oscillatory information to other microstate generator regions.

Milz et al. (2016) found microstate class A to be associated with visual and class B with verbal processing when manipulating the modality of information processing during eyes-closed EEG recordings (all parameters were affected: *duration*, *occurrence* and *coverage*). This appears to be in contradiction with Britz et al.'s (2010) findings. However, EEG microstates are predominantly driven by intra-cortical sources in the alpha band frequency (Milz et al., 2017) and alpha is known to have

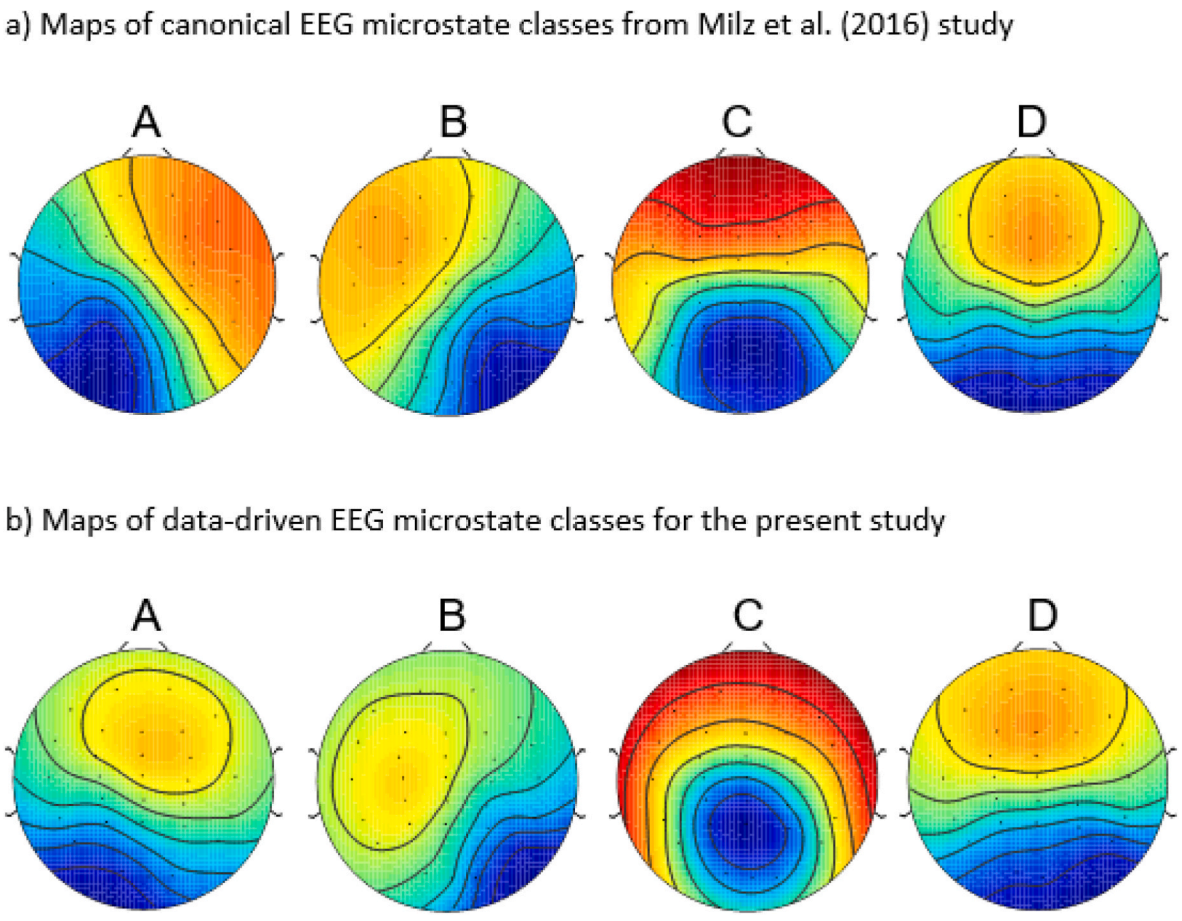


Fig. 1. The topographical maps for canonical EEG microstate classes A, B, C, and D: a) Milz et al. (2016) maps from the KeyPy template library (reduced from 64 to 30 channels and rendered with the EEGLAB topoplot function); and b) data-driven maps from the present study's recordings computed in KeyPy as the population average over all subjects and conditions, and classified into the canonical classes (only recordings that had no rejected channels and lasting at least 120 s after pre-processing were included, $n = 95$). These two sets of topographical maps explained 69% and 73% of the variance of the momentary maps at the GFP peaks of the present recordings, respectively. For data-driven maps derived at the subject- and condition-level (not shown in the figure), explained variances were higher, with 78% for the GFP peaks and 72% for all time-frames. Explained variances were calculated using equation (12) of Pascual-Marqui et al. (1995).

inhibitory effects on modality-specific processing (O'Gorman et al., 2013). Conversely, negative BOLD activity might be related to inhibitory neuronal activity (Sten et al., 2017). Taken together, the associations of classes A and B with *negative* BOLD signal in phonological and visual processing brain areas respectively in the Britz et al. (2010) study, as well as Pascual-Marqui et al.'s (2014) findings, suggest that Milz et al.'s observation of the predominance of microstate class A during visualization might reflect inhibiting action on left-hemispheric language processing areas, and class B during verbalization might reflect inhibiting action on right-hemispheric visuo-spatial processing areas. These cross-modality inhibitory effects of alpha frequency have been described previously (e.g. Cooper et al., 2003).

However, contrary to this reasoning, Seitzman et al. (2017) observed significant increases in the *coverage* and *occurrence* of class B microstate when participants transitioned from an eyes-closed to eyes-open state thereby increasing visual input (and supposedly decreasing verbal processing associated with mind-wandering during the eyes-closed resting state), suggesting class B to be associated with visual rather than verbal processing.

The functional significance of class C microstate is also unclear. Milz et al. (2016) and Seitzman et al. (2017) reported its decreased *occurrence* during a task relative to a resting condition. Seitzman et al. (2017) proposed that class C microstate might reflect the activity of the Default Mode Network (DMN) known to decrease during goal-oriented cognitive processing. This is in contradiction with the suggestion by Britz et al. (2010) that class C reflects the activity of the cognitive control/error monitoring networks. Their study also observed no associations of the four classes with the DMN, despite an ICA component representing the DMN being present in the BOLD signal.

Similarly, the findings in relation to class D microstate are mixed. Whereas Seitzman et al. (2017) have reported increased *duration* and *occurrence* of class D during a serial subtraction task compared to rest either with eyes closed or open, Milz et al. (2016) reported increased *duration* and *occurrence* of microstate D during rest compared to task conditions. Thus, it is presently unclear whether class D reflects attention needed to sustain a goal-directed cognitive activity, as suggested by Seitzman et al. (2017), or a rapid switching and reorientation of attention that could occur more frequently during 'rest' than during a sustained attention condition, as suggested by Milz et al. (2016). Given that Britz et al. (2010) observed the association of class D with the *negative* BOLD in the dorsal attentional network, class D topography could reflect an inhibition of this network in line with Milz et al.'s finding of the increased presence of class D microstate during rest, which is known to be associated with attenuated activity of the top-down attention network.

The four canonical microstates have the potential to be developed into a clinical biomarker, as their parameters show differences between neuropsychiatric populations and healthy controls (review, Khanna et al., 2015). For example, first-episode medication-naïve patients with schizophrenia show decreased average *duration* of class B (Lehmann et al., 2005; Irisawa et al., 2006; Nishida et al., 2013) and class D (Lehmann et al., 2005; Kikuchi et al., 2007; Nishida et al., 2013), and increased occurrence of class A (Lehmann et al., 2005) and class C (Lehmann et al., 2005; Kikuchi et al., 2007; Nishida et al., 2013) relative to healthy controls. Tomescu et al. (2014) reported decreased *occurrence* of class D and increased *occurrence* of class C in adolescents with the 22q11 deletion syndrome, a condition associated with a 30-fold increased risk of developing schizophrenia. Adults with mood and anxiety disorders have been reported to have a longer *duration* of class C compared with healthy controls (Al Zoubi et al., 2019).

Canonical EEG microstates may offer potential as biomarkers for early detection and characterization of dementia. Decreased average *duration* of all four canonical microstates has been reported in people with mild to moderate Alzheimer's disease, but not with age-related cognitive decline (Dierks et al., 1997; Strik et al., 1997; Stevens and Kircher, 1998). Also, frontotemporal dementia patients show decreased class C *duration* compared to patients with Alzheimer's disease and healthy controls (Nishida et al., 2013).

Microstate syntax also holds promise as a clinical tool, having been shown to differentiate clinical populations from each other and from healthy populations. Lehmann et al. (2005) reported the reversal of the class sequence in first-episode medication-naïve patients with schizophrenia (A→D→C→A) compared with healthy controls (A→C→D→A), Al Zouby et al. (2019) found adults with mood and anxiety disorders to have higher A→D and B→C transition probabilities alongside a lower B→D transition probability compared with healthy controls, whereas Nishida et al. (2013) observed Alzheimer's disease patients to have a reversal of the predominant microstate pairwise transition (D→C) compared with healthy controls (C→D).

Given the growing interest in EEG microstate applications in a clinical context, with some promising evidence for the four canonical ones, yet conflicting findings in relation to their functional significance, the main aim of the present study was to further clarify the functional significance of the four canonical microstates by manipulating the information processing modality, verbal vs. visual, during eyes-closed EEG in a healthy population. A further novel aim was to investigate the short-term (within a single session) test-retest reliability of microstate parameters and their pairwise transitions; establishing test-retest reliability is an important step towards enabling a rigorous use of any measure either for basic neuroscience research or as a clinical biomarker.

2. Methods

2.1. Participants

A total of 24 participants were recruited via university circular emails and local on-line forums. The inclusion criteria were right-handedness and age between 18 and 65 years. The exclusion criteria included a history of mental health problems, drug and/or alcohol abuse, and neurodevelopmental or neurodegenerative disorders. Three participants were excluded due to left-handedness as ascertained by the Edinburgh Handedness Inventory (Oldfield, 1971), and a further participant due to a pre-existing neurological condition.

The final sample analysed contained 20 participants (16 males, mean age = 31.5, standard deviation = 12.5).

Ethics approval was gained from King's College London Ethics Approval Board (Ref: HR-16/17–4092) and informed consent was gained from the participants prior to participation.

2.2. Experimental paradigm

To clarify the functional significance of EEG microstates, we designed a paradigm enabling a better control over verbal and visual processing within and between individuals than the protocol used by Milz et al. (2016), which would have engaged complex cognitive processes in their verbal condition (i.e. executive function and semantic processing to define a word) and possibly affective processing in their object-visual condition (i.e. in response to viewing pictures of animals). Additionally, we aimed to minimize the differences between the verbal and visual conditions as much as possible, setting them to diverge *only* on the processing modality. To achieve this, we opted for the silent verbalization of the word 'square' vs. visualization of a 2-dimensional square. Two pilot experiments were conducted to develop the paradigm; their description and results are reported in the Supplementary Materials, Appendix I.

The final experimental paradigm consisted of two identical runs of two repetitions of the same condition sequence: *mind-wandering* -> *verbalization* -> *visualization* (i.e. Run 1 & 2: *mind-wandering* -> *verbalization* -> *visualization* (Repetition 1) -> *mind-wandering* -> *verbalization* -> *visualization* (Repetition 2)).

For the *mind-wandering* condition, the participants were instructed to

relax and allow their minds to wander naturally. For the *verbalization* and *visualization* conditions, the participants were asked to repeat silently the word 'square' or to visualize a square, respectively, at a self-paced rate of approximately 2 s (without silent counting). Each condition was 3 min long, with a total duration of 18 min for each run. Participants had their eyes closed throughout the run. All participants were given a practice run of a full condition sequence lasting 1 min.

2.3. Subjective experience ratings

After each experimental run, the participants had a period of rest (approx. 5–10 min), during which they were debriefed about their performance overall and their overall levels of alertness and focus during verbalization and visualization, as well as asked about their predominant thought form (verbal vs. visual) during the mind-wandering conditions. The participants were then asked to rate different aspects of their experience using visual analogue scales (VASs), separately for the first and second half of each run. The VASs rated on a scale of 0–100%: i) the level of alertness (with 0 being 'very drowsy/not alert at all' and 100 being 'fully alert'); ii) the amount of mind-wandering experienced during the verbalization and visualization conditions (with 0 being 'little to no mind wandering' and 100 being 'constant mind wandering'); and iii) the amount of effort needed to stay focused on each task (with 0 being 'little to no effort/the task was very easy' and 100 being 'a lot of effort/the task was very difficult'). After Run 2 only, two further VASs were administered to rate: i) the amount of 'interference' experienced during verbalization and visualization (i.e. to what extent a visualization of square occurred spontaneously during verbalization condition and vice versa, with 0 being 'no interference' and 100 being 'constant interference'); ii) a predominance of either verbal or visual content during the mind-wandering condition (with 0 being 'entirely verbal' and 100 being 'entirely visual').

2.4. Experimental paradigm presentation

The paradigm (written instructions; an example of a square outline for the visualization condition; and VASs for the subjective performance ratings) was programmed in OpenSesame (Mathôt et al., 2012) and presented on a 24" Dell SE2416H6 LCD monitor. Auditory instructions before the start of each condition were relayed using a prerecorded voice presented through the PC speakers.

2.5. EEG data collection & analysis

2.5.1. EEG data recording

EEG data were recorded using a 40-channel Neuroscan Quikcap system (Compumedics, USA). A total of 36 channels were recorded including 4 electrodes for the EOG signal and two references. The EEG electrodes were arrayed according to the standard 10–20 EEG setup. The VEOG electrodes were positioned above and below the left eye, the HEOG electrodes at the outer canthi. The online reference electrode was set to A2, and the ground electrode was at AFz. The data were low-pass filtered at 250 Hz and sampled at 1 kHz.

2.5.2. EEG data preprocessing

The EEG data were inspected in Brainstorm (Tadel et al., 2011), imported into EEGLAB (Delorme and Makeig, 2004), and segmented into twelve ~ 3-min epochs (i.e. 2 runs x 2 repetitions x 3 conditions = 12 epochs). Each EEG epoch was re-referenced to the common average and band-pass filtered (1–20 Hz, 4th-order Butterworth filter). Eye-movement artifacts were removed by regressing the EEG channels with the EOG (Croft and Barry, 2002). Obvious (mostly movement and swallowing) artifacts were cut out (ranging between 0 and 11 in number per epoch) and noisy electrodes rejected (0–3 per epoch) before the application of independent component analysis (ICA) on the remaining channels (27–30 per epoch) and time-frames (93,000–185,000 per epoch). For each ICA component, the time-course, 2D skull map of inverse weights, and power spectrum were inspected, and the percentage variance of the EEG signal explained by each component was compared to the explained variance of the EOG signal. Standard criteria were used for component rejection based on their spatial profile and temporal-frequency composition. The number of retained ICA components varied between recording epochs (range 7–26) but had little effect on the derived micro-state outcomes, as also observed by others (Dinov and Leech, 2017). Finally, the EEG data were down-sampled to 250 Hz.

2.5.3. EEG microstate computation

The EEG Analysis Toolbox KeyPy (Milz, 2015; Milz et al., 2016) assigns each EEG time-frame to a microstate class based on its 2D spatial configuration. Microstate classes can be pre-defined, or calculated from the data using a clustering algorithm. For the present study, both strategies were used. First, as a quality test for our recordings, four data-driven maps were generated by running (200 repetitions of) KeyPy's modified k-means clustering algorithm, and assigning them the labels A, B, C and D based on their correlation with the canonical class maps of Milz et al. (2016) (see Fig. 1b for the data-driven population class maps). Only the time-frames at the local global-field-power (GFP) peaks were used to minimize the effect of noise, with low-power intermediate time-frames being interpolated by KeyPy. As most of the GFP peaks occur at the crests and troughs of the alpha waves, the average interval between peaks was about 50 ms (50.2 ± 3.4 over all epoch).

In a second stage, KeyPy was rerun skipping the clustering phase and assigning directly the peaks of the GFP function to one of the four canonical maps of Milz et al. (2016) (as with the data-driven maps, the Milz et al. maps were first assigned to the GFP peaks and interpolated in between).

The method of assigning the GFP peaks to the four canonical classes using Milz et al. maps of Fig. 1a, instead of using the data-driven population maps of Fig. 1b, was necessitated by the aim of assessing test-retest reliability and the paradigm design. If average maps over different runs and repetitions (and by extension over different conditions and participants) would have been used, then the map parameters of the first presentation would have been affected by the maps of all subsequent repetitions, potentially confounding the measurement of reliability.

Fig. 2 shows the spatial correlations between Milz et al. (2016) and our data-driven population maps. We note that our class A correlates

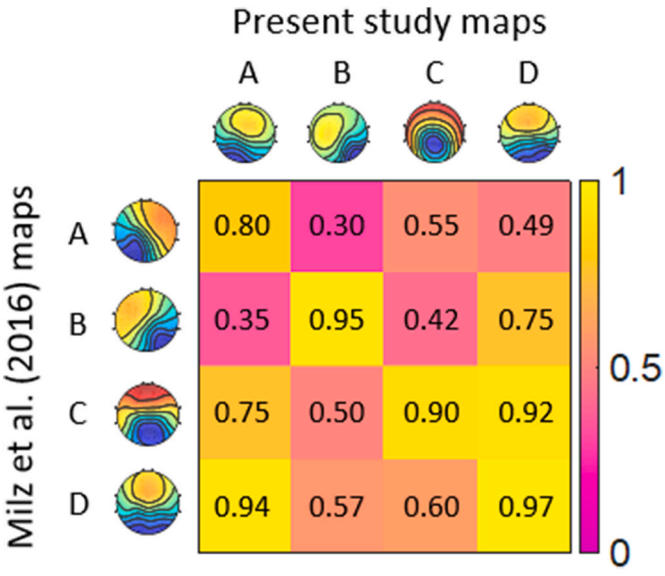


Fig. 2. Pearson correlation coefficients between Milz et al. (2016) population maps and the data-driven population maps from our sample.

more strongly with the Milz et al. class D (r = 0.94) than A (r = 0.80). It is not uncommon to have high correlations between the maps of different classes in a dataset; for example, the Milz et al. maps C and D have a correlation of 0.86. The reason for the great similarity between the Milz et al. map D and our map A is that some of our maps labelled as D look slightly anti-clockwise rotated. However, the labelling across the set of four classes maximizes the match of the data-driven maps from our study with the Milz et al. canonical population maps, yielding the best match across all 24 permutations.

The parameters of each microstate class: mean duration (milliseconds), number of occurrences (per second), and time coverage (%) as defined in previous research (e.g. Koenig et al., 1999), were calculated from the canonical maps. The microstate classes and their parameters were derived for each repetition (epoch) of each condition in each run separately per participant. Due to the rejection of 'noisy' segments during pre-processing, the final length of each epoch varied, with the highest durations of group mean epochs being 166–169 s. Paired *t*-tests comparing mean epoch durations were not significant, apart from the first repetitions of the verbalization condition of two runs (mean group duration of 165 s for Run 1 and 173 s for Run 2, $p < .04$), confirming there were no systematic differences in epoch lengths that could confound the test-retest reliability analysis.

2.5.4. Conditional transition probabilities between microstates

Transitions between microstates, an aspect of microstate *syntax* (Lehmann et al., 2005), have been looked at in a number of different ways in previous research, including as counts or frequencies of 12 (or 16) pairs (depending on whether self-transitions, e.g. A→A, are included or not), or as conditional probabilities of transitioning to a target state starting from a given source state. Whereas much of the early literature reported counts or frequencies of transition pairs, which has the advantage of being able to directly compare the opposite pairs (e.g. A→B vs. B→A) (Lehmann et al., 2005; Schlegel et al., 2012; Nishida et al., 2013; Spring et al., 2018), conditional transition probabilities (as used in the theory of stochastic processes) have commonly been utilized in both the early (Wackermann et al., 1993) and more recent studies (Brodbeck et al., 2012; Gärtner et al., 2015; Tomescu et al., 2014, 2015, 2018; Liu et al., 2017; Bréchet et al., 2018).

Seitzman et al. (2017), who employed cognitive manipulation to probe the functional associations of the microstates as per our aims, used conditional transition probabilities and presented matrices for sequences with and without self-transitions (their Fig. 4, left and right columns, respectively), whilst von Wegner et al. (2017) examined the symmetry of the transition matrix.

We used conditional transition probabilities between the 12 pairs of microstates (excluding self-transitions) to make our findings on the functional significance and test-retest reliability of microstate syntax relevant to/comparable with the most recent research. In brief, all self-transitions were removed from the microstate sequences for each epoch in order to produce a sequence of transitions between different microstates. There were no systematic (i.e. statistically significant) differences in the total number of microstate transitions per epoch, with the mean number of transitions per epoch ranging between 1899 and 2051.

For each epoch, the *conditional transition probabilities* were calculated as:

$P_{X\to Y} = n_{XY}/n_X$, where n_{XY} is the number of XY pairs in the sequence, and n_X the number of Xs.

To test whether the observed transition probabilities were significantly different from what would be expected from the microstates' occurrences, the *observed* numbers of pairs XY:

$$n_{XY} = n_X P_{X\to Y}$$

were compared to those *expected* from the microstate occurrences under the null hypothesis that the distinct microstate occurring next is independent of the current one.

If P_Z denotes the relative occurrence of microstate Z, then the expected transition probability $P^*_{X\to Y}$ and the *expected* number n^*_{XY} of pairs XY are (modifying the formula from Lehmann et al., 2005):

$$P^*_{X\to Y} = \frac{P_Y}{1 - P_X}$$

$$n^*_{XY} = n_x P^*_{X\to Y}$$

Since the generation of pairs XY can be considered as a random process with a binomial distribution, where $p = P^*_{X\to Y}$ is the probability that X is followed by Y and $q = 1\text{-}p$ the probability that X is followed by a state different from Y (and X), the standard deviation s^*_{XY} of the number of pairs XY is given by:

$$s^*_{XY} = \sqrt{n_X pq}$$

The numbers of transitions are typically large (given the sequence length and only 12 different transition pairs), so that, under the null hypothesis, the binomial distribution approaches a standard normal distribution after the transformation of:

n_{XY} to $(n_{XY} - n^*_{XY})/s^*_{XY}$

The transformed values were averaged across all 20 subjects within the same condition.

Averaged values greater than 1.96 or less than −1.96 were considered significant at $p < .05$ level.

2.6. Statistical data analysis strategy

Statistical analyses were performed using IBM SPSS software version 24. Prior to statistical testing, all variables were assessed for normality of the distribution by comparing their skewness with the standard error (Coolican, 2009). Where variables were found to be non-normally distributed, box plots were used to identify the outliers, whose data were then removed from the analysis. A total of 16 data points across 14 variables were removed as outliers (no single participant had more than 2 data points removed).

The test-retest reliability of the microstate class parameters and their conditional transition probabilities (referred to as 'transition probabilities' thereafter) within and between the runs was estimated using the Intraclass Correlation Coefficients (ICC, Landis and Koch, 1977). The ICCs were calculated using a 2-way mixed-effects model with an absolute agreement definition for a single measure (Koo and Li, 2016; Aldridge et al., 2017).

We applied Koo and Li (2016) guidelines for ICC estimates, based on the 95% confidence interval, with values less than 0.5, between 0.5 and 0.74, between 0.75 and 0.9, and greater than 0.90 being indicative of poor, moderate, good, and excellent reliability, respectively. For the clarity of reporting, we have rounded up the ICCs to one decimal point (e.g. ICCs between 0.71 and 0.74 were rounded up to 0.7 and between 0.75 and 0.79 up to 0.8), which gives the following ranges: < 0.5 = poor; 0.5–0.7 = moderate; 0.8–0.9 = good, > 0.9 (ICC of 0.9 with CI between 0.8 and 1) = excellent reliability. The ICCs above 0.7 were considered to indicate adequate test-retest reliability since they account for at least half of the variance (Post, 2016; Streiner and Norman, 2008).

Since most ICCs were 0.7 or less, with most confidence intervals (CIs) being wide, with the lower CI margin falling into the poor range (<0.5), indicating less than adequate test-retest reliability, we entered each repetition of each condition separately into the analysis of variance models when investigating the functional significance of four microstate classes. This resulted in a 4x3x2x2 repeated-measures analysis of variance (rmANOVA) model with class (A, B, C, and D), condition (mind-wandering, verbalization, visualization), run (run 1 and run 2), and repetition (2 repetitions per run) for each parameter: *duration*, *occurrence*, and *coverage*. Main effects and interactions were followed up with lower level rmANOVAs and paired *t*-tests as appropriate. *P* values for

contrasts investigating the main effects were adjusted for multiple comparisons as follows: main effect of class $p < .008$ (6 contrasts), main effect of condition $p < .02$ (3 contrasts), main effects of run and repetition $p < .03$ (2 contrasts). The interaction effects and their contrasts are also reported for the p values at $p < .1$ for which the partial eta squared (an estimate of the effect size) ranged from medium to large ($\eta_p^2 > 0.13$), as they were considered to be meaningful trends given the small sample size.

The same analysis strategy was employed for the transition probabilities with the rmANOVAs followed by lower-level ANOVA to further investigate the main effects and interactions.

Pearson correlations were used to investigate the relationships of the microstate parameters and the transition probabilities with the subjective ratings using the VASs.

3. Results

3.1. Test-retest reliability of alpha power

Since the EEG microstates are mainly driven by the alpha frequency, as a reference, we first report the relative alpha power and its test-retest reliability.

The fraction of power located in the alpha band (8–12 Hz; 1–20 Hz band-pass filtered EEGs) declined gradually across the six conditions of each run (Mind-wandering1, Verbalization1, Visualization1, Mind-wandering2, Verbalization2, Visualization2), with the steepest decline during the first three conditions (Run 1: 0.53, 0.50, 0.46, 0.47, 0.47 and 0.46; Run 2: 0.56, 0.55, 0.51, 0.50, 0.49 and 0.50). Accordingly, mean alpha power over the four repetitions was slightly higher during mind-wandering (0.52 ± 0.18) and verbalization (0.50 ± 0.18) than during visualization (0.48 ± 0.20). In contrast, test-retest reliability was higher during mental tasks, especially visualization, than during mind-wandering. It was also generally higher during the first two conditions of each run, when the participants were more alert (see below), except for the visualization condition, which had excellent test-retest reliability between two repetitions of the same run for both runs (see Table S1).

Relative alpha power was strongly correlated with the microstate duration (Pearson correlations significant at the 0.01 level in 8 of the 12 conditions), with the correlations being stronger for classes A ($r = 0.60$) and B ($r = 0.64$) than for C ($r = 0.22$) and D ($r = 0.44$).

3.2. Test-retest reliability of canonical microstate classes

Supplementary Table S2 presents ICCs with confidence intervals for pairwise contrasts of four condition repetitions over two runs for each microstate class parameter.

For the microstate classes, class A had adequate reliability (ICC $\geq$0.7) on all parameters for the mind-wandering condition. For the verbalization condition, it had adequate reliability for *duration,* with mixed reliability for other parameters. For the visualization condition, it had good to excellent reliability (ICC range 0.7–0.9) for *coverage* and *occurrence*.

Class B had a rather poor test-retest reliability (ICC range 0.4–0.6) for *coverage* and *occurrence* in the mind-wandering condition, achieving adequate reliability (ICC range 0.7–0.9) only for *duration*. It did better in the verbalization and visualization conditions, with good to excellent reliability (ICC range 0.7–0.9) on all parameters, with the only exception of moderate reliability for *duration* (ICC of 0.6) during the second repetitions of the visualization condition in both runs.

Class C had the most mixed test-retest reliability for all conditions on all parameters, except for *occurrence* during mind-wandering for which it showed adequate reliability.

Class D appears to be the most reliable out of the four classes for *duration* and *coverage* (ICC range 0.7–0.9), but showed quite poor reliability for *occurrence* across the conditions.

The first repetitions of two runs had the highest test-retest reliability across the microstate class parameters, followed by two repetitions of Run 1 with a few exceptions (i.e. class B for *coverage* and *occurrence* in mind-wandering where two repetitions of Run 2 performed better).

Overall, the test-retest reliability of the microstate class parameters appears to differ depending on the condition. The verbalization condition produced the highest but mostly moderate-to-good reliability for all four classes for *duration*, but this was not the case for *coverage* and *occurrence* for which the visualization condition performed most stably. *Duration* appears to be the most reliable parameter overall, particularly for classes A and B during mind-wandering and verbalization, and for class D in all conditions.

3.3. Functional significance of canonical microstates classes

3.3.1. Analysis of variance for each parameter with run and repetition as fixed factors

Table 1 presents the means and standard deviations for four classes in each repetition of the three conditions over two runs for each parameter. It also reports the statistics for the significant main effects and interactions of the 4x3x2x2 rmANOVAs with class (A, B, C, and D), condition (mind-wandering, verbalization, visualization), run (run 1 and run 2), and repetition (2 repetitions per run), as well as the results of the lower-level rmANOVAs following up main effects and interactions. Fig. S1 displays the estimated marginal means from the rmANOVAs for each class in three conditions per run per parameter. Below we summarize the main effects and interactions across the parameters, and present the statistics for the lower-level pairwise contrasts not included in Table 1, where relevant.

As can be seen from Table 1 and Fig. S1, the means for class D were the highest in all conditions for all the parameters, followed by class C, with classes A and B being comparable. Consequently, the main effect of class was significant for all the parameters, with the lower level contrasts indicating that class D was significantly longer in *duration* and higher in *occurrence* and *coverage* than the other classes, with class C being significantly longer in *duration* and higher in *occurrence* and *coverage* than either class A or class B.

The main effect of condition was significant for *duration* and *occurrence*. For *duration*, the main effects of lower level rmANOVAs contrasting the conditions pairwise were not significant (at the corrected p value). For *occurrence*, lower level rmANOVAs showed the main effect of condition to be due to the differences between mind-wandering and visualization, with paired t-tests indicating greater *occurrence* of class A (trend) during the first repetition [$t_{19} = -1.9$, $p = .08$] and of class B during the first [$t_{19} = -2.09$, $p = .008$] and second [$t_{19} = -2.5$, $p = .02$] repetitions of visualization than mind-wandering in Run 1 only.

There were no significant class by condition interactions in the full rmANOVA models, indicating that none of the parameters of four canonical classes differed systematically depending on the information processing modality across both runs and four repetitions. The only significant class by condition interaction was observed for mind-wandering vs. verbalization [$F_{3,\ 42} = 3.8$, $p = .02$, $\eta_p^2 = 0.21$] when exploring the main effect of condition for *duration* using lower levels rmANOVAs, with the paired t-tests having revealed a significantly longer mean *duration* of class D during the first repetition of mind-wandering than verbalization [$t_{19} = 2.5$, $p = .02$] in Run 1.

There was a significant main effect of run for *occurrence*, driven by a significant main effect for class C and a trend (at corrected p value of .01) for class A. Class x run interaction was also significant, with lower-level rmANOVAs indicating that the interaction was driven by the differences between classes A and C vs. D, due to the significantly greater *occurrence* of class A during the first repetition of verbalization [$t_{18} = 3.7$, $p = .002$] and visualization [$t_{19} = 2.4$, $p = .02$] and of class C during the second repetition of mind-wandering [$t_{19} = 2.3$, $p = .03$] in Run 1 than Run 2.

Class x run interaction was also significant for *coverage*, but none of the lower-level rmANOVAs contrasting the classes pairwise survived the correction for multiple comparisons ($p = .008$). The uncorrected p values

Table 1
Descriptive and inferential statistics for four classes for each parameter: *duration* (ms), *coverage* (%), and *occurrence* (per second).

Parameter	Class	Run1_Con_Rep	Mean (SD)	Run2_Con_Rep	Mean (SD)	Significant main effects and interactions	Pairwise contrasts for significant effects
Duration	**A**	Run1_MW1	72.3 (13.6)	Run2_MW1	71.9 (12.2)	**Class**: $F_{1.4,\ 14.1}$[a] = 10.4, $\boldsymbol{p = .003}$, η_p^2 = .51	**Class:**
		Run1_MW2	71.3 (14.2)	Run2_MW2	72.7 (14.3)	**Condition**: $F_{2,\ 20} = 3.8$, $\boldsymbol{p = .04}$, $\eta_p^2 = .27$	**A vs. D**: $F_{1,\ 15} = 23.1$, $\boldsymbol{p < .001}$, $\eta_p^2 = .61$
		Run1_Ver1	72.6 (12.7)	Run2_Ver1	74.5 (16.3)		**B vs. D**: $F_{1,\ 15} = 19.8$, $\boldsymbol{p < .001}$, $\eta_p^2 = .57$
		Run1_Ver2	69.0 (9.6)	Run2_Ver2	73.1 (12.4)		**C vs. D**: $F_{1,\ 15} = 10.9$, $\boldsymbol{p = .005}$, $\eta_p^2 = .42$
		Run1_Vis1	70.3 (9.4)	Run2_Vis1	70.2 (8.3)		**A vs. B**: $F_{1,\ 14} = 1.04$, $p = .09$, $\eta_p^2 = .19$
		Run1_Vis2	70.0 (10.6)	Run2_Vis2	69.9 (12.2)		**A vs. C**: $F_{1,\ 13} = 4.43$, $\boldsymbol{p = .05}$, $\eta_p^2 = .25$
	B	Run1_MW1	73.4 (13.3)	Run2_MW1	74.7 (14.7)		**B vs. C**: $F_{1,\ 13} = 10.5$, $\boldsymbol{p = .006}$, $\eta_p^2 = .45$
		Run1_MW2	70.7 (12.6)	Run2_MW2	74.5 (14.2)		**Condition:**
		Run1_Ver1	73.8 (13.0)	Run2_Ver1	78.5 (17.7)		**MW vs. Ver:** $F_{1,\ 14} = 1.7$, $p = .22$, $\eta_p^2 = .11$
		Run1_Ver2	71.3 (14.5)	Run2_Ver2	71.4 (13.7)		
		Run1_Vis1	73.1 (10.4)	Run2_Vis1	73.8 (14.4)		**MW vs. Vis**: $F_{1,\ 11} = 1.9$, $p = .19$, $\eta_p^2 = .15$
		Run1_Vis2	73.8 (12.3)	Run2_Vis2	73.8 (15.4)		**Ver vs. Vis**: $F_{1,\ 12} = 0.2$, $p = .65$, $\eta_p^2 = .02$
	C	Run1_MW1	84.2 (20.5)	Run2_MW1	86.2 (21.1)		
		Run1_MW2	86.7 (25.2)	Run2_MW2	81.1 (16.8)		
		Run1_Ver1	84.0 (16.8)	Run2_Ver1	86.3 (18.2)		
		Run1_Ver2	86.4 (18.3)	Run2_Ver2	82.6 (20.4)		
		Run1_Vis1	81.8 (28.9)	Run2_Vis1	109.7 (35.4)		
		Run1_Vis2	84.4 (18.0)	Run2_Vis2	83.2 (18.0)		
	D	Run1_MW1	104.8 (25.6)	Run2_MW1	110.5 (38.2)		
		Run1_MW2	101.9 (33.7)	Run2_MW2	109.9 (40.2)		
		Run1_Ver1	95.8 (24.8)	Run2_Ver1	109.1 (37.8)		
		Run1_Ver2	101.4 (29.8)	Run2_Ver2	102.6 (29.9)		
		Run1_Vis1	98.4 (28.9)	Run2_Vis1	109.7 (35.4)		
		Run1_Vis2	100.5 (32.6)	Run2_Vis2	106.1 (36.7)		

Parameter	Class	Run1_Con_Rep	Mean (SD)	Run2_Con_Rep	Mean (SD)	Significant main effects and interactions	Pairwise contrasts for significant effects
Coverage	**A**	Run1_MW1	18.4 (7.1)	Run2_MW1	17.5 (7.8)	**Class:** $F_{2.0,\ 38.3}$[a] = 14.7, $\boldsymbol{p < .001}$, $\eta_p^2 = .45$	**Class:**
		Run1_MW2	18.2 (6.9)	Run2_MW2	18.3 (6.4)	**Class*Run:** $F_{2.2,\ 42.4}$[a] = 4.2, $\boldsymbol{p = .02}$, η_p^2 = .19	**A vs. D**: $F_{1,\ 19} = 29.8$, $\boldsymbol{p < .001}$, $\eta_p^2 = .61$
		Run1_Ver1	19.2 (6.4)	Run2_Ver1	17.5 (6.1)		**B vs. D**: $F_{1,\ 19} = 22.6$, $\boldsymbol{p < .001}$, $\eta_p^2 = .54$
		Run1_Ver2	17.2 (4.8)	Run2_Ver2	19.2 (6.0)		**C vs. D**: $F_{1,\ 19} = 4.6$, $\boldsymbol{p = .05}$, $\eta_p^2 = .19$
		Run1_Vis1	19.3 (5.5)	Run2_Vis1	17.4 (6.4)		**A vs. B**: $F_{1,\ 19} = 1.1$, $p = .31$, $\eta_p^2 = .05$
		Run1_Vis2	17.5 (6.2)	Run2_Vis2	17.2 (6.2)		**A vs. C**: $F_{1,\ 19} = 11.3$, $\boldsymbol{p = .003}$, $\eta_p^2 = .37$
	B	Run1_MW1	18.4 (6.2)	Run2_MW1	18.9 (7.9)		**B vs. C**: $F_{1,\ 19} = 8.7$, $\boldsymbol{p = .008}$, $\eta_p^2 = .31$
		Run1_MW2	17.8 (6.9)	Run2_MW2	19.6 (6.2)		**Class*Run:**
		Run1_Ver1	20.4 (7.1)	Run2_Ver1	20.5 (6.9)		**A vs. D**: $F_{1,\ 19} = 6.2$, $p = .02$, $\eta_p^2 = .25$
		Run1_Ver2	18.2 (6.9)	Run2_Ver2	18.2 (7.3)		**B vs. D**: $F_{1,\ 19} = 3.1$, $p = .09$, $\eta_p^2 = .14$
		Run1_Vis1	20.6 (5.0)	Run2_Vis1	19.2 (6.8)		**C vs. D**: $F_{1,\ 19} = 7.8$, $p = .01$, $\eta_p^2 = .29$
		Run1_Vis2	20.9 (6.6)	Run2_Vis2	19.7 (7.2)		**A vs. B**: $F_{1,\ 19} = 0.3$, $p = .57$, $\eta_p^2 = .02$
	C	Run1_MW1	26.6 (9.6)	Run2_MW1	26.2 (10.6)		**A vs. C**: $F_{1,\ 19} = 2.3$, $p = .15$, $\eta_p^2 = .11$
		Run1_MW2	29.7 (13.8)	Run2_MW2	24.6 (10.9)		**B vs. C**: $F_{1,\ 19} = 2.2$, $p = .16$, $\eta_p^2 = .10$
		Run1_Ver1	27.6 (8.6)	Run2_Ver1	26.9 (11.1)		
		Run1_Ver2	29.4 (10.1)	Run2_Ver2	26.8 (10.9)		
		Run1_Vis1	27.0 (9.2)	Run2_Vis1	25.5 (9.8)		
		Run1_Vis2	28.1 (10.1)	Run2_Vis2	26.5 (12.3)		
	D	Run1_MW1	36.5 (11.3)	Run2_MW1	37.4 (14.2)		
		Run1_MW2	34.3 (12.4)	Run2_MW2	37.5 (13.1)		
		Run1_Ver1	32.7 (10.3)	Run2_Ver1	35.1 (11.9)		
		Run1_Ver2	35.2 (13.0)	Run2_Ver2	35.9 (13.2)		
		Run1_Vis1	33.1 (11.6)	Run2_Vis1	37.8 (13.6)		
		Run1_Vis2	33.4 (14.6)	Run2_Vis2	36.6 (13.0)		

Parameter	Class	Run1_Con_Rep	Mean (SD)	Run2_Con_Rep	Mean (SD)	Significant main effects and interactions	Pairwise contrasts for significant effects
Occurrence	**A**	Run1_MW1	2.5 (0.8)	Run2_MW1	2.4 (0.9)	**Class:** $F_{3,\ 54} = 12.2$, $\boldsymbol{p < .001}$, $\eta_p^2 = .40$	**Class:**
		Run1_MW2	2.5 (0.7)	Run2_MW2	2.5 (0.7)	**Condition**: $F_{1.5,\ 27.1}$[a] = 3.4, $\boldsymbol{p = .06}$, η_p^2 = .16	**A vs. D**: $F_{1,\ 18} = 29.7$, $\boldsymbol{p < .001}$, $\eta_p^2 = .62$
		Run1_Ver1	2.6 (0.8)	Run2_Ver1	2.3 (0.7)		**B vs. D**: $F_{1,\ 19} = 19.3$, $\boldsymbol{p < .001}$, $\eta_p^2 = .51$

(*continued on next page*)

Table 1 (*continued*)

Parameter	Class	Run1_Con_Rep	Mean (SD)	Run2_Con_Rep	Mean (SD)	Significant main effects and interactions	Pairwise contrasts for significant effects
		Run1_Ver2	2.5 (0.6)	Run2_Ver2	2.6 (0.7)	**Run:** $F_{1, 18} = 5.9$, $\boldsymbol{p = .03}$, $\eta_p^2 = .25$	**C vs. D**: $F_{1, 19} = 1.5$, $p = .23$, $\eta_p^2 = .08$
		Run1_Vis1	2.7 (0.7)	Run2_Vis1	2.5 (0.8)	**Class*Run:** $F_{3, 54} = 4.1$, $\boldsymbol{p = .01}$, $\eta_p^2 = .19$	**A vs. B**: $F_{1, 18} = 1.6$, $p = .23$, $\eta_p^2 = .08$
		Run1_Vis2	2.5 (0.8)	Run2_Vis2	2.4 (0.8)		**A vs. C**: $F_{1, 18} = 11.9$, $\boldsymbol{p = .003}$, $\eta_p^2 = .4$
	B	Run1_MW1	2.5 (0.7)	Run2_MW1	2.5 (0.8)		**B vs. C**: $F_{1, 19} = 8.8$, $\boldsymbol{p = .008}$, $\eta_p^2 = .32$
		Run1_MW2	2.5 (0.8)	Run2_MW2	2.6 (0.7)		**Condition:**
		Run1_Ver1	2.7 (0.7)	Run2_Ver1	2.6 (0.7)		**MW vs. Ver**: $F_{1, 18} = 2.6$, $p = .12$, $\eta_p^2 = .11$
		Run1_Ver2	2.5 (0.6)	Run2_Ver2	2.5 (0.8)		
		Run1_Vis1	2.8 (0.6)	Run2_Vis1	2.6 (0.8)		**MW vs. Vis**: $F_{1, 18} = 5.6$, $\boldsymbol{p = .02}$, $\eta_p^2 = .23$
		Run1_Vis2	2.8 (0.7)	Run2_Vis2	2.7 (0.8)		
	C	Run1_MW1	3.1 (0.7)	Run2_MW1	3.0 (0.9)		**Ver vs. Vis**: $F_{1, 148} = 1.5$, $p = .24$, $\eta_p^2 = .08$
		Run1_MW2	3.3 (1.0)	Run2_MW2	3.0 (1.0)		
		Run1_Ver1	3.2 (0.7)	Run2_Ver1	3.1 (1.0)		**Run:**
		Run1_Ver2	3.4 (0.8)	Run2_Ver2	3.2 (0.9)		**A**: $F_{1, 18} = 5.4$, $\boldsymbol{p = .03}$, $\eta_p^2 = .23$
		Run1_Vis1	3.3 (0.8)	Run2_Vis1	3.1 (0.9)		**B**: $F_{1, 18} = .7$, $p = .41$, $\eta_p^2 = .04$
		Run1_Vis2	3.3 (0.8)	Run2_Vis2	3.0 (1.1)		**C**: $F_{1, 19} = 7.4$, $\boldsymbol{p = .01}$, $\eta_p^2 = .28$
	D	Run1_MW1	3.5 (0.6)	Run2_MW1	3.4 (0.7)		**D**: $F_{1, 98} = .4$, $p = .54$, $\eta_p^2 = .02$
		Run1_MW2	3.3 (0.5)	Run2_MW2	3.4 (0.6)		**Class*Run:**
		Run1_Ver1	3.4 (0.6)	Run2_Ver1	3.3 (0.5)		**A vs. D**: $F_{1, 18} = 6.3$, $\boldsymbol{p = .02}$, $\eta_p^2 = .23$
		Run1_Ver2	3.4 (0.8)	Run2_Ver2	3.4 (0.7)		**B vs. D**: $F_{1, 19} = 1.3$, $p = .28$, $\eta_p^2 = .06$
		Run1_Vis1	3.3 (0.6)	Run2_Vis1	3.4 (0.6)		**C vs. D**: $F_{1, 19} = 6.6$, $\boldsymbol{p = .02}$, $\eta_p^2 = .26$
		Run1_Vis2	3.2 (0.9)	Run2_Vis2	3.5 (0.7)		**A vs. B**: $F_{1, 18} = .7$, $p = .42$, $\eta_p^2 = .2$
							A vs. C: $F_{1, 18} = 3.1$, $p = .1$, $\eta_p^2 = .15$
							B vs. C: $F_{1, 19} = 1.9$, $p = .18$, $\eta_p^2 = .1$

Abbreviations: Con = Condition; MW = Mind-Wandering; Rep = Repetition; Ver = Verbalization; Vis = Visualization.
[a] Greenhouse-Geisser Corrected Statistics.

indicated the interaction was due to the differences between classes A and D, with paired *t*-tests showing a greater *coverage* of class A during the first repetition of verbalization [$t_{19} = 2.2$, $p = .04$] and visualization [$t_{19} = 2.1$, $p = .04$] in Run 1 than Run 2, but a greater coverage of class D for the second repetition of verbalization [$t_{19} = -1.8$, $p = .08$] and the first repetition of visualization [$t_{19} = -2.8$, $p = .01$] in Run 2 than Run 1. The *t*-test contrasts were significant at uncorrected *p*-value only, but, given the small sample size, we note them here as trend indicators of test-retest reliability of microstate class parameters.

There were no main effects or interactions for repetition.

3.3.2. *Analysis of variance for the parameters averaged across runs and repetitions*

Given that there were no main effects of repetition in the 4x3x2x2 rmANOVAs, we conducted 4x3 rmANOVAs with class (A, B, C, D) and condition (mind-wandering, verbalization, and visualization) as within-subject factors, with the parameters averaged across four repetitions over two runs to allow for more power in detecting the differences, if any, in the microstate class parameters as a result of the experimental manipulation in processing modality (i.e. verbal vs visual). Significant main effects and interactions were followed up with lower level rmANOVAs and paired *t*-tests.

As can be seen from Table 2, there were significant main effects of class for all three parameters, and of condition for *duration* and *occurrence*. As before, the effect of class was due to higher means for all parameters for Class D, followed by class C, with no differences between classes A and B. The main effects of condition were due to class D having longer *duration* in mind-wandering than verbalization conditions [$t_{19} = 2.3$, $p = .03$], and class B having higher *occurrence* in visualization than either in mind-wandering [$t_{19} = -3.3$, $p = .004$] or verbalization [$t_{19} = -2.9$, $p = .008$] conditions.

The only observed class by condition interaction was for *duration* at a trend level ($p < .1$), which was due to class D having longer *duration* in mind-wandering than verbalization and visualization conditions compared to class B [$F_{2, 30} = 2.8$, $p = .08$, $\eta_p^2 = 0.16$] and class C [$F_{2, 30} = 3.1$, $p = .06$, $\eta_p^2 = 0.17$]. Lower level rmANOVAs have also revealed a significant class by condition interaction for class B with higher *occurrence* in visualization than in either mind-wandering or verbalization conditions compared to class A [$F_{2, 36} = 4.5$, $p = .02$, $\eta_p^2 = 0.20$].

3.4. *Conditional transition probabilities between EEG microstates*

3.4.1. *Test-retest reliability of conditional transition probabilities*

Table S3 presents the ICCs and their CIs for the transition probabilities for each pair of the four repetitions of two runs. Overall, the ICCs were even more mixed than those observed for the microstate parameters ranging from 0.3 to 0.9, with a substantial proportion of the ICCs falling below adequate test-retest reliability (<0.7). The ICCs across four repetitions were the highest for D→A during mind-wandering and A→B during verbalization and visualization (ICCs range 0.7–0.9), with the

Table 2
Table of Mean and SD among EGG microstates averages for *duration*, *coverage*, and *occurrence* by condition.

Parameter	Class	Mind-wandering Mean (SD)	Verbalization Mean (SD)	Visualization Mean (SD)	Significant main effects and interactions
Duration	**A**	69.5 (9.8)	72.3 (12.1)	67.9 (6.6)	**Class**: $F_{3, 30} = 10.4$, $p = .003$, $\eta_p^2 = .51$
	B	73.3 (12.3)	72.6 (12.9)	70.7 (8.1)	**Condition**: $F_{2, 20} = 3.8$, $p = .04$, $\eta_p^2 = .27$
	C	80.3 (13.3)	84.8 (16.9)	79.8 (12.7)	**Class x Condition:**
	D	106.8 (31.2)	102.2 (28.1)	99.8 (26.3)	$F_{6, 60} = 5.9$, $p = .08$, $\eta_p^2 = .17$
Coverage	**A**	18.1 (6.3)	18.3 (5.2)	17.8 (5.5)	**Class**: $F_{3, 30} = 14.4$, $p < .0001$, $\eta_p^2 = .44$
	B	18.7 (5.7)	19.3 (6.4)	20.1 (5.8)	
	C	26.7 (10.0)	27.7 (8.9)	26.8 (9.0)	
	D	36.4 (11.1)	34.7 (10.7)	35.2 (12.1)	
Occurrence	**A**	2.5 (0.7)	2.4 (0.5)	2.5 (0.7)	**Class**: $F_{3, 54} = 12.2$, $p < .0001$, $\eta_p^2 = .40$
	B	2.5 (0.6)	2.6 (0.7)	2.7 (0.6)	**Condition**: $F_{2, 36} = 3.4$, $p = .04$, $\eta_p^2 = .16$
	C	3.1 (0.9)	3.2 (0.8)	3.2 (0.8)	
	D	3.4 (0.5)	3.4 (0.6)	3.4 (0.6)	

remaining ICCs being at 0.7 or less for all transition probabilities across three conditions. Low test-retest reliability was observed for A→C during visualization and C→B during mind-wandering with the ICCs ranging between 0.4 and 0.7 (overall ICC of 0.5), as well as C→D during mind-wandering with all but one ICCs being between 0.3 and 0.6 (overall ICC of 0.5). The worst performing were the probabilities of transitioning from D→B during mind-wandering (ICCs range 0.3–0.6) and visualization (all but one ICC ranging between 0.3 and 0.6).

3.4.2. *Functional significance of conditional transition probabilities*

Fig. 3 presents the transition matrices of the mean transition probability for each pair of the four microstates for the four repetitions of each condition across two runs. (Fig. S2 presents the means, standard deviations, and the diagrams of the transition probabilities by their relative likelihood, i.e. mean probability). As can be seen from Fig. 3, the probabilities of transitioning from C→D and D→C were the highest across the conditions in all repetitions, except for D→C for the second repetition of mind-wandering in Run 2. These conditional transition probabilities were also significantly greater than those predicted from the occurrences (under the null hypothesis that the identity of the next microstate is independent of the current one, at $p < .05$) for 10 of the 12 C→D values and for 9 of the 12 D→C values. Next highest probabilities were for A→D and B→D transitions, particularly in verbalization and visualization conditions. The mean probability of the C→B transition remained consistently one of the lowest in all repetitions of all conditions, particularly during visualization, with 2 of the 12 C→B probabilities being significantly lower than predicted from the null hypothesis (at $p < .05$). Overall, the transitions matrices were asymmetrical, with directional effects for all but the A→B transitions.

Noteworthy, the canonical microstate maps were not uncorrelated (in terms of their topographical similarity), with maps C and D being correlated most strongly ($r = 0.87$), followed by B vs. C ($r = 0.68$), B vs. D ($r = 0.63$), A vs. D ($r = 0.6$), A vs. C ($r = 0.53$), and with A vs. B being correlated only modestly ($r = 0.2$). Although transition probabilities do not depend on the correlations between the maps *a priori*, in a noisy EEG signal such correlations may lead to a misallocation of momentary topographic maps at GFP peaks to a 'wrong' class, leading to spurious microstate transitions where there were none.

3.4.3. *Asymmetry of transition probabilities*

To test the significance of the differences in the asymmetry of transition probabilities between the conditions, we conducted separate (6 in total, i.e. A→B vs. B→A, A→C vs. C→A, A→D vs. D→A, B→C vs. C→B, B→D vs. D→B, and C→D vs. D→C) 2x3x2x2 rmANOVAs with a transition (e.g. A→B vs. B→A), condition (mind-wandering, verbalization, and visualization), run (1 and 2) and repetition (1 and 2). Table 3 presents the main results of these analyses.

As can be seen in Table 3, the only observed significant transition by

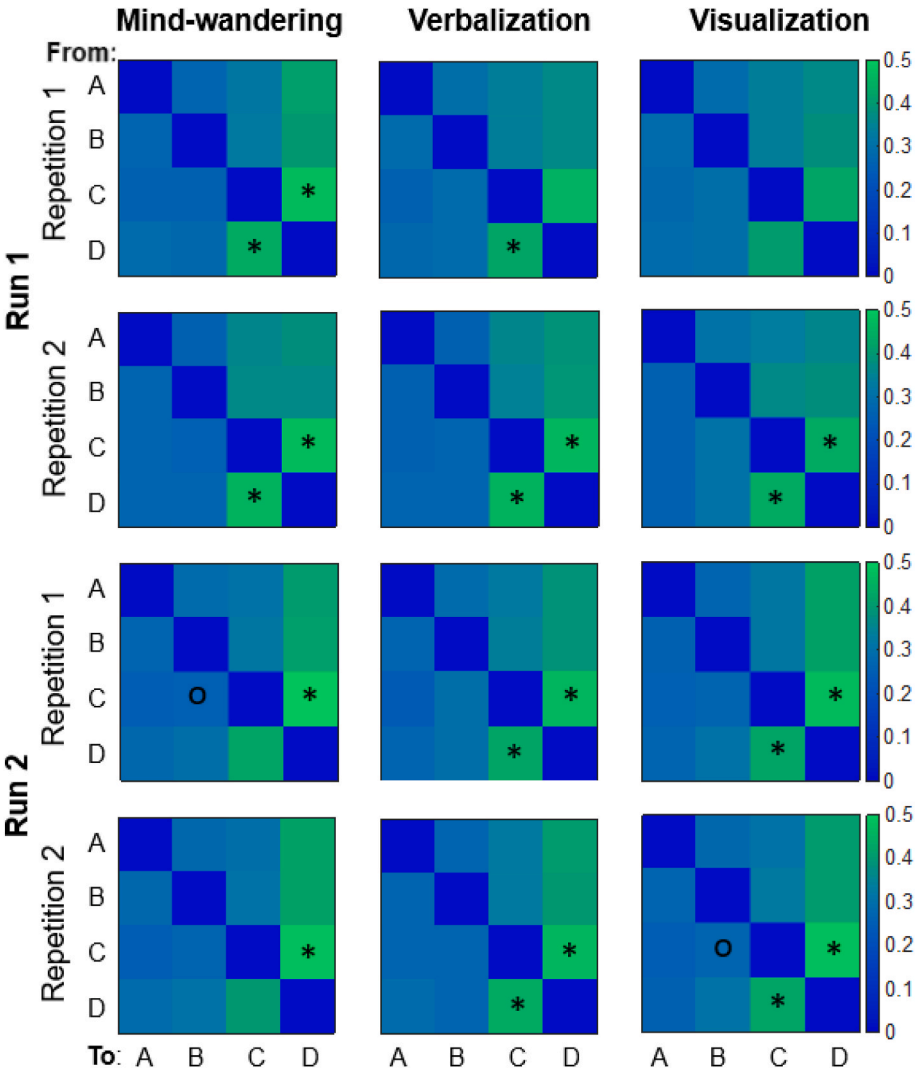


Fig. 3. Transition probability matrices for the two runs of two repetitions of the three conditions. Probabilities were averaged over the 20 subjects. Values significantly greater than those expected from the microstate occurrences are marked by an asterisk, those significantly lower by a circle.

Table 3
Significant main effects and interactions of the repeated measures analysis of variance examining the asymmetry of transition probabilities, and the data patterns that explain them.

Transition direction	Significant main effects and interactions	Data pattern
A→B vs. B→A	-	
A→C vs. C→A	**Transition:** $F_{1, 19} = 10.7$, $p = .004$, $\eta_p^2 = .36$ **Run:** $F_{1, 19} = 7.1$, $p = .01$, $\eta_p^2 = .27$	A→C > C→A Both pairs: Run1 > Run2
A→D vs. D→A	**Transition:** $F_{1, 19} = 26.3$, $p < .0001$, $\eta_p^2 = .58$ **Run:** $F_{1, 19} = 3.9$, $p = .06$, $\eta_p^2 = .17$ **Transition x Run:** $F_{1, 19} = 7.6$, $p = .01$, $\eta_p^2 = .29$	A→D > C→D A→D: Run2 > Run1
B→C vs. C→B	**Transition:** $F_{1, 19} = 7.6$, $p = .01$, $\eta_p^2 = .29$ **Run:** $F_{1, 19} = 5.8$, $p = .03$, $\eta_p^2 = .24$ **Transition x Rep:** $F_{1, 19} = 3.6$, $p = .07$, $\eta_p^2 = .16$	B→C > C→B[a] Both pairs: Rep1&2 Run1
B→D vs. D→B	**Transition:** $F_{1, 18} = 17.8$, $p = .001$, $\eta_p^2 = .50$ **Run:** $F_{1, 18} = 6.3$, $p = .02$, $\eta_p^2 = .26$ **Condition x Rep:** $F_{2, 36} = 3.2$, $p = .05$, $\eta_p^2 = .15$	B→D > D→B Both pairs: Run2 > Run1 B→D: Rep1&2 Run2 > Rep1&2 Run1 in MW and Vis vs. Ver
C→D vs. D→C	**Transition x Condition:** $F_{2, 18} = 3.2$, $p = .05$, $\eta_p^2 = .27$ **Direction x Run:** $F_{1, 19} = 9.1$, $p = .007$, $\eta_p^2 = .32$ **Condition x Run:** $F_{1, 19} = 9.1$, $p = .007$, $\eta_p^2 = .32$	MW vs. Ver: C→D > D→C MW vs. Vis: C→D > D→C Run 2 Vis: C→D > D→C Run2 > Run1

Abbreviations: MW = Mind-Wandering; Rep = Repetition; Ver = Verbalization; Vis = Visualization.

[a] C→B transition probability was significantly lower than predicted from the occurrences of microstates B and C in mind-wandering repetition 1 of run 2 and visualization repetition 2 of run 2.

condition interaction was for the probability of C→D transition being significantly higher than for D→C during mind-wandering compared to verbalization. There were a number of direction and condition by run or repetition interaction effects, suggesting a lack of reproducibility of transition probabilities from run to run and repetition to repetition.

3.4.4. *Transition probability to and from a class*

We also examined whether transition probabilities *to* and *from* each class significantly differed by condition to further probe functional significance of the pairwise transitions. Table 4 presents the main findings from the 3x3x2x2 rmANOVAs with transition-*to*-class (e.g. B→A vs. C→A vs. D→A) or transition-*from*-class (e.g. A→B vs. A→C vs. A→D), condition (mind-wandering, verbalization, visualization), run (1 and 2), and repetition (1 and 2) (a total of 8 rmANOVA models).

As can be seen from Table 4, there were a number of significant transition by condition interactions for the transitions *to* or *from* a class, but none held up across runs and repetitions in the lower level contrasts. There were also a number of significant transition by condition by run and transition by condition by repetition interactions (see Table 4 for data patterns).

Overall, there were many more significant main effects and interactions by run for transition probabilities than for microstate class parameters, and a number of significant interactions by repetition (not observed for class parameters), which is in line with an overall poorer test-retest reliability (reproducibility) of transition probabilities than microstate class parameters. This was particularly true for the transitions *from* class C.

Table 4
Significant main effects and interactions of the repeated measures analysis of variance examining the differences in transition probabilities *to* and *from* each class, and the data patterns that explain them.

Transition *to* class	Significant main effects and interactions	Data pattern
Class A (B→A vs. C→A vs. D→A)	**Transition x Condition x Repetition:** $F_{4, 16} = 4.7$, $p = .01$, $\eta_p^2 = .54$	B→A > C→A in Ver Rep1 Run1&2
Class B (A→B vs. C→B vs. D→B)	**Condition:** $F_{2, 17} = 3.5$, $p = .05$, $\eta_p^2 = .29$ **Transition x Condition:** $F_{2, 17} = 3.5$, $p = .05$, $\eta_p^2 = .29$ **Condition x Run:** $F_{2, 17} = 4.3$, $p = .03$, $\eta_p^2 = .34$	All to class B: Vis > MW or Ver Run1 A→B > C→B in MW vs. Ver Run2
Class C (A→C vs. B→C vs. D→C)	**Transition:** $F_{2, 38} = 23.9$, $p < .0001$, $\eta_p^2 = .73$ **Run:** $F_{1, 19} = 4.8$, $p = .04$, $\eta_p^2 = .20$ **Transition x Condition x Repetition:** $F_{4, 16} = 3.2$, $p = .04$, $\eta_p^2 = .44$	D→C > A→C or B→C A→C and B→C: Run 2 > Run 1 in MW and Vis A→C: Ver > MW Rep1 Run2 and Vis Rep2 Run2 D→C: Ver > Vis Rep2 Run2
Class D (A→D vs. B→D vs. C→D)	**Transition:** $F_{2, 38} = 25.2$, $p < .0001$, $\eta_p^2 = .57$ **Run:** $F_{1, 19} = 8.0$, $p = .01$, $\eta_p^2 = .30$ **Transition x Condition:** $F_{4, 76} = 2.4$, $p = .06$, $\eta_p^2 = .11$	C→D > A→D or B→D C→D: MW > Vis Run1
Transition *from* class[a]		
Class A (A→B vs. A→C vs. A→D)	**Transition:** $F_{2, 38} = 5.3$, $p = .01$, $\eta_p^2 = .21$ **Repetition:** $F_{1, 19} = 5.6$, $p = .03$, $\eta_p^2 = .23$ **Transition x Run:** $F_{2, 38} = 4.5$, $p = .02$, $\eta_p^2 = .19$ **Condition x Run:** $F_{2, 38} = 3.4$, $p = .04$, $\eta_p^2 = .15$ **Transition x Condition x Run:** $F_{4, 76} = 2.5$, $p = .05$, $\eta_p^2 = .12$	A→D > A→B, particularly in Run 2 A→D: MW > Vis Rep1 Run1
Class B (B→A vs. B→C vs. B→D)	**Transition:** $F_{2, 38} = 6.3$, $p = .004$, $\eta_p^2 = .25$ **Transition x Run:** $F_{2, 38} = 6.3$, $p = .004$, $\eta_p^2 = .25$ **Condition x Repetition:** $F_{2, 38} = 4.4$, $p = .02$, $\eta_p^2 = .19$	B→C and B→D > B→A B→C: Ver > MW Rep1 Run2
Class C (C→A vs. C→B vs. C→D)	**Transition:** $F_{2, 38} = 28.5$, $p < .0001$, $\eta_p^2 = .60$ **Condition:** $F_{2, 38} = 6.5$, $p = .004$, $\eta_p^2 = .27$ **Run:** $F_{1, 19} = 5.9$, $p = .03$, $\eta_p^2 = .24$ **Repetition:** $F_{1, 19} = 6.2$, $p = .02$, $\eta_p^2 = .25$ **Transition x Condition:** $F_{4, 76} = 3.8$, $p = .007$, $\eta_p^2 = .17$ **Transition x Run:** $F_{2, 38} = 4.2$, $p = .02$, $\eta_p^2 = .18$ **Condition x Run:** $F_{2, 38} = 6.1$, $p = .005$, $\eta_p^2 = .24$ **Condition x Repetition:** $F_{2, 38} = 8.0$, $p = .001$, $\eta_p^2 = .30$ **Run x Repetition:** $F_{1, 19} = 7.3$, $p = .01$, $\eta_p^2 = .28$ **Transition x Condition x Run:** $F_{4, 76} = 5.6$, $p < .0001$, $\eta_p^2 = .23$	C→D > C→A or C→B C→D: MW > Vis Run 1 C→B: Vis > MW Rep1&2 Run1; Vis > Ver Rep2 Run 1

(*continued on next page*)

Table 4 (*continued*)

Transition *to* class	Significant main effects and interactions	Data pattern
	Condition x Run x Repetition: $F_{2, 38} = 7.1$, $p = .002$, $\eta_p^2 = .27$	
Class D (D→A vs. D→B vs. D→C)	**Transition:** $F_{2, 36} = 17.6$, $p < .0001$, $\eta_p^2 = .50$	D→C > D→A or D→B

Abbreviations: MW = Mind-Wandering; Rep = Repetition; Ver = Verbalization; Vis = Visualization.

[a] Note that the transition probabilities *from* a class add up to one since for each EEG microstate class X, transition probabilities $p_{X,Y}$ *from* X to the other microstate classes Y add up to one, i.e., $\sum_{Y \neq X} p_{X,Y} = 1$.

3.5. Subjective experience ratings

3.5.1. Alertness

The participants rated themselves as being more alert at the start of each run, as well as at the start rather than the end of the second sequence of each run, which is somewhat surprising considering that one would expect the participants to be drowsier during mind-wandering (start of the run/sequence) than visualization (end of the run/sequence) (see Table 5). These subjective ratings are in agreement with higher relative alpha power during the start of each run and the start of each sequence (see subsection 3.1).

The 2x2x2 rmANOVA with run (1 and 2), sequence (1 and 2), and the sequence segment (start vs. end) as the within-subject factors confirmed the significant main effects of sequence [$F_{1, 18} = 19.2$, $p < .0001$, $\eta_p^2 = 0.52$] and the sequence segment [$F_{1, 18} = 47.1$, $p < .0001$, $\eta_p^2 = 0.72$].

There were a number of significant correlations between subjective alertness ratings with class parameters. (Please note that given the way the alertness ratings were sampled (sequence start vs. end), we correlated the class parameters for verbalization condition with both sequence start and end alertness ratings, only with the alertness ratings at the start of the sequence for mind-wandering condition, and only with the alertness ratings at the end of the sequence for visualization condition.) The alertness ratings were significantly positively correlated with class A *duration* during mind-wandering [$r_{17} = 0.55$, $p = .02$] and verbalization [$r_{19} = 0.47$, $p = .05$] for the start, and during verbalization [$r_{19} = 0.47$, $p = .05$] and visualization [$r_{19} = 0.51$, $p = .03$] for the end of sequence 1 in run 1. There were also negative correlations with class D *occurrence* during mind-wandering [$r_{19} = -0.53$, $p = .02$] for the start of sequence 1 in run 1, as well with class C *occurrence* during verbalization [$r_{19} = -0.52$, $p = .02$] and visualization [$r_{19} = -0.45$, $p = .05$], and class C *coverage* during verbalization [$r_{19} = -0.51$, $p = .03$] for the end of sequence 1 in run 1.

There were also a number of significant correlations with transition probabilities, including: positive correlations with B→A probability [$r_{19} = 0.46$, $p = .05$] and D→A [$r_{19} = 0.65$, $p = .003$] and negative correlations with A→C [$r_{19} = -0.52$, $p = .03$] and D->C [$r_{19} = -0.59$, $p = .008$] probabilities during verbalization (start of sequence), as well as a positive correlation with D→A [$r_{19} = 0.49$, $p = .04$] and a negative correlation with D→C [$r_{19} = -0.46$, $p = .05$] probabilities during visualization (end of sequence) for sequence 1 in run 2. There was also a positive correlation with C→A probability during visualization (end of sequence) [$r_{19} = 0.54$, p = .02], and a negative correlation with D→A probability during verbalization for both the start [$r_{19} = -0.46$, p = .05] and the end [$r_{19} = -0.43$, p = .07] of sequence 2 in run 2 (note that a positive correlation of D→A probability with alertness ratings during verbalization was observed for the start of sequence 1 in run 1, see above).

3.5.2. Effortfulness

As can be seen from Table 6, presenting the means, SDs and the range (minimum-maximum values) for the VAS ratings of effortfulness, there were substantial individual differences in how effortful the participants have found both conditions. A 2x2x2 rmANOVA with condition (verbalization vs. visualization), run (1 and 2) and repetition (1 and 2) as within-subject factors showed a significant main effect of condition [$F_{1, 18} = 14.9$, $p = .001$, $\eta_p^2 = 0.45$], with effortfulness being rated as higher during visualization than verbalization in all repetitions. There was also a significant condition by repetition interaction [$F_{1, 18} = 6.9$, $p = .02$, $\eta_p^2 = 0.23$], with significantly higher ratings during verbalization for repetition 1 in run 2 than run 1.

Higher effort was associated with shorter class B *duration* during visualization for repetition 2 in run 1 [$r_{19} = -0.57$, p = .01].

In run 2, for repetition 1 of the verbalization condition higher effort was associated with shorter class D *duration* [$r_{19} = -0.48$, p = .04] and higher class C *occurrence* [$r_{19} = 0.54$, $p = .02$] and *coverage* [$r_{19} = 0.51$, $p = .02$], as well as higher transition probabilities for A→C [$r_{19} = 0.46$, $p = .05$], B→C [$r_{19} = 0.58$, $p = .009$] and a trend for D→C [$r_{19} = 0.43$, $p = .07$], as would be expected from the positive correlation with higher class C *occurrence*. Additionally, for repetition 1 of the visualization condition higher effort was associated with shorter class A [$r_{19} = -0.58$, $p = .009$] and class B [$r_{19} = -0.46$, $p = .05$] *duration*, as well as higher class C *occurrence* [$r_{19} = 0.52$, $p = .02$] and *coverage* [$r_{19} = 0.47$, $p = .04$]. For repetition 2 of the visualization condition, higher effort correlated with shorter class A *duration* [$r_{19} = -0.49$, $p = .05$], as well as lower class A [$r_{19} = -0.48$, $p = .04$] and class B [$r_{19} = -0.52$, $p = .02$] *coverage*. Finally, higher effort correlated with higher probability of D→C transition [$r_{19} = 0.56$, $p = .01$] during repetition 1 and lower probability of A→B and B→A transitions during both repetition 1 [A→B: $r_{19} = -0.57$, $p = .02$; B→A: $r_{19} = -0.48$, $p = .04$] and repetition 2 [A→B: $r_{19} = -0.63$, $p = .004$; B→A: $r_{19} = -0.56$, $p = .01$] of the visualization condition.

3.5.3. Interference

The participants varied substantially in the degree of spontaneous mind-wandering during the verbalization and visualization conditions, as well as spontaneous visual and verbal interference during the verbal and visual conditions respectively (see Table 7 for the means, SDs, and ranges for both ratings). Higher degree of spontaneous mind-wandering correlated with higher effortfulness ratings, indicating a greater subjective effort to stay on task in the presence of mind-wandering.

Table 5
Means, SDs and minimum-maximum values (range) for the subjective ratings of alertness using the VAS with 0% 'being not at all alert' and 100% 'being extremely alert'. The ratings were sampled after each run with the participants asked to rate their alertness at the start and the end of each sequence of the run: sequence 1 (Mind-wandering1->Verbalization1->Visualization1) and sequence 2 (Mind-wandering2->Verbalization2->Visualization2).

Alertness rating	Sequence start Mean (SD) [Range]	Sequence end Mean (SD) [Range]
Run1_Sequence1	84.3 (10.6) [70.4–100]	57.4 (21.4) [9.2–83.6]
Run1_Sequence2	74.2 (16.1) [39–95.2]	47.1 (21.7) [11.8–78.8]
Run2_Sequence1	82.0 (14.5) [42.8–100]	60.5 (22.4) [15.8–91.4]
Run2_Sequence2	72.7 (22.2) [20–100]	47.6 (24.4) [3.6–89.0]

Table 6
Means, SDs and minimum-maximum values (range) for the subjective ratings of effortfulness using the VAS with 0% 'being not at all effortful' and 100% 'being extremely effortful'. The ratings were sampled after each run with the participants asked to rate how effortful they found each repetition of the verbalization and visualization conditions.

Effortfulness rating	Verbalization condition Mean (SD) [Range]	Visualization condition Mean (SD) [Range]
Run1_Repetition1	45.6 (24.1) [12.4–90.6]	63.1 (23.1) [6.6–100]
Run1_Repetition2	48.5 (23.2) [12.2–85.6]	70.3 (25.5) [22.6–100]
Run2_Repetition1	50.5 (28.2) [4.2–94]	64.3 (25.3) [18–100]
Run2_Repetition2	49.8 (29.4) [5–97]	67.4 (28.2) [2.8–100]

Table 7
Means, SDs and ranges for the subjective interference ratings of: i) spontaneous mind-wandering during verbalization and visualization conditions; and ii) spontaneous verbalization (i.e. repeating 'square') during the visualization condition and spontaneous visualization (i.e. visualizing 'square') during the verbalization condition. To rate this, two separate Visual Analogue Scales (VASs) were administered after run 2: VAS 1 rated the amount of interference experienced during verbalization and visualization (i.e. to what extent a visualization of square occurred spontaneously during verbalization condition and vice versa) with 0% 'no interference' and 100% 'constant interference', and VAS 2 rated a predominance of either verbal or visual content during the mind-wandering condition with 0% 'entirely verbal' and 100% 'entirely visual'.

VAS	Interference type	Run_Repetition	Condition	
			Verbalization Mean (SD) [Range]	**Visualization** Mean (SD) [Range]
VAS 1	**Mind-wandering**	Run1_Repetition1	46.9 (34.7) [0–100]	55.3 (26.4) [4–100]
		Run1_Repetition2	53.5 (25.8) [9.8–99.2]	62.0 (22.9) [21.6–100]
		Run2_Repetition1	42.1 (30.6) [5–100]	49.9 (26.1) [6.2–100]
		Run2_Repetition2	45.1 (28.7) [6.4–100]	59.5 (29.7) [1–100]
VAS 2	**Verbalizing 'square'**	Overall	–	50.4 (28.7) [0–94.6]
	Visualizing 'square'	Overall	42.6 (32.7) [0–98]	–

A 2x2x2 rmANOVA with condition (verbalization vs. visualization), run (1 and 2) and repetition (1 and 2) showed a trend for the main effect of condition [$F_{1, 18} = 4.0$, $p = .06$, $\eta_p^2 = 0.18$], with a higher degree of spontaneous mind-wandering indicated during visualization than verbalization.

Higher degree of spontaneous mind-wandering was correlated with higher class D *occurrence* during verbalization [$r_{19} = 0.67$, $p = .002$] and visualization [$r_{19} = 0.63$, $p = .004$], and higher probability of A→D [$r_{19} = 0.51$, $p = .03$] and C→D [$r_{19} = 0.47$, $p = .04$] transitions during visualization in repetition 1 of run 1, as well as higher class D *occurrence* [$r_{19} = 0.50$, $p = .03$] during verbalization in repetition 2 of run 1.

No significant correlations were observed for the spontaneous visual and verbal interference ratings during the verbal and visual conditions with either the average class parameters or transition probabilities.

4. Discussion

Given a rising interest in the EEG microstates and their potential application as biomarkers of healthy and pathological neural dynamics, in this study we: i) further probed into the functional significance of the four canonical EEG microstate classes; and ii) investigated within-session test-retest reliability of their parameters and conditional transition probabilities in healthy participants. The results of the study have important implications for the assumptions of stationarity (replicability) of the parameters characterizing the four canonical microstate classes most used in previous research, and particularly for their transition probabilities. Our findings also highlight that it is premature to assign clear functional associations to the four canonical EEG microstate classes.

4.1. Test-retest reliability of microstate parameters and transition probabilities

In terms of within-session test-retest reliability, we found the parameters of microstate classes to be more stable overall than their conditional transition probabilities. The microstate *duration* was the most reproducible parameter (except for classes A and B during visualization), whereas *occurrence* was least reliable. The poor reliability of the transition probabilities is likely to be the consequence of the poor reliability of *occurrence*, as the transition probabilities are driven by the microstate occurrences.

The run had an effect on the parameters as evidenced by a few significant class by run interactions, with seemingly random fluctuations in the parameters from run to run (between-run repetitions). Although there were no significant class by condition by run (or repetition) interactions for any of the parameters, and the overall (across all repetitions) ICCs were acceptable, it is nevertheless recommended that the ICCs are checked in the future studies before making the decision to average the parameters across the repetitions in the same run/session. Alternatively, the "global" microstate class parameters could be derived across the sessions/recordings, as was implemented by the only other study to assess test-retest reliability of microstate parameters (Khanna et al., 2014), which did so across three sessions spaced over 48 h apart of resting-state eyes-closed EEG (no processing modality manipulation was involved), and observed high reliability for the microstate maps derived over all sessions/recordings, but low/mixed reliability for the maps derived separately for each session or recording, as is also the case in our study. Simply increasing the number of runs/trials within a single session, an approach most commonly taken towards minimizing the intra- and inter-subjective 'noise', is likely to only generate more randomness in the parameters due to the fluctuating brain and mental states, as evidenced by the drop in the alpha power from the start to the end of each run and lower subjective ratings of alertness towards each condition sequence and each run. A particular caution should be taken by studies adopting test-retest designs, e.g. clinical trials using EEG microstate parameters as an outcome measure of an intervention or longitudinal studies of psychopathological, neurodevelopmental, or neurodegenerative conditions. Instead of increasing the number of trials/runs, the experimental designs should allow for shorter epochs with the resting breaks between the repetitions of the same condition or a sequence of conditions.

The within-session test-retest reliability of the transition probabilities, reported here for the first time to the best of our knowledge, was truly problematic with poor reproducibility across repetitions of the same condition for most transition pairs. Some ICCs were as low as 0.3, with the majority of the ICCs failing to reach the threshold for the adequate test-retest reliability (>0.7). The least stable across repetitions were the transition probabilities from class C, particularly C→B and C→D during mind-wandering, as well as for D→B during the mind-wandering and visualization conditions. These findings call for great caution in using transition probabilities as biomarkers of clinical populations for which the intra- and inter-subject 'noise' tends to be higher than in the general population.

However, a number of observed associations of transition probabilities with the subjective ratings of alertness and effortfulness suggest the sensitivity of transition probabilities to mental states. This, in turn, might prove useful in developing pairwise transition probabilites and longer microstate sequences as clinical biomarkers; for example, as predictors of treatment response to a psychotherapeutic intervention, if a greater understanding of the relationship between the experiential (subjective) dynamics and the neural dynamics as indexed by microstate syntax can be reached.

4.2. Functional significance of EEG microstates

In terms of functional significance, as concluded by Milz et al. (2016) and as confirmed by our findings, it would be erroneous to assign distinct functional associations to either of the four canonical microstates, as all four appear during different information processing modalities (verbal vs. visual), as well as mind-wandering. Despite the predicted dominance of classes A and B during visual and verbal information processing respectively, based on the findings by Milz et al. (2016), class D had the longest *duration* as well as highest *coverage* and *occurrence*, followed by class C, regardless of the condition.

Nevertheless, there are a few clues from the results of our study as to the associations of the four canonical classes with cognitive processes. A number of findings appear to link class D with mind-wandering. This is suggested by the longer class D *duration* during mind-wandering than verbalization (repetition 1, run 1), as well as a trend for class by condition interaction for the averaged parameters, with class D *duration* being longer than classes B and C during mind-wandering than during verbalization or visualization. Further, higher alertness ratings were associated with lower class D *occurrence* during mind-wandering (run 1), whilst a higher degree of spontaneous mind-wandering was associated with a greater class D *occurrence* during verbalization and visualization (run 1), as well as with higher A→D and C→D transition probabilities. Considering that the mean ratings for the spontaneous mind-wandering during verbalization and visualization were approx. 50% of the time, the dominance of class D on all parameters in all conditions is most likely explained by its association with spontaneous mind-wandering.

The putative association of class D with mind-wandering and/or drowsiness in Milz et al. (2016) and in our study, however, is in conflict with Seitzman et al. (2017) who have reported increased *duration* and *occurrence* of class D during a serial subtraction task compared to rest either with eyes closed or open. Furthermore, patients with schizophrenia show decreased class D *duration* (Koenig et al., 1999; Lehmann et al., 2005; Kikuchi et al., 2007; Nishida et al., 2013) relative to healthy controls, which is contrary to what would be expected if, indeed, class D microstate was associated with mind-wandering and the associated activation of the DMN known to be hyperactive in schizophrenia (Whitfield-Gabrieli and Ford, 2012). However, it should be noted that the topography of class D in schizophrenia patients is of altered configuration as compared to the canonical class D microstate seen in the healthy controls across the studies (e.g. Koenig et al., 1999). Likewise, maps C and D of Seitzman et al. (2017) seem to resemble more maps D and C, respectively, of Milz et al. (2016) (see upper row of Fig. 1 for the latter).

The indirect support for the class D association with the DMN comes from the study comparing EEG microstate maps during wakefulness and three stages of non-REM sleep (Brodbeck et al., 2012), which observed a reduction of class D *coverage* in deeper sleep. The fMRI studies of resting state networks during deeper non-REM sleep found reduced functional connectivity between the main nodes of the DMN (Horovitz et al., 2009; Sämann et al., 2011). Further research is required to elucidate the functional significance of class D microstate and the neural networks/-dynamics associated with it, particularly as its parameters have been found to be the most replicable ones in our study and it has been shown to differentiate clinical populations (Nishida et al., 2013; Tomescu et al., 2015), making it a candidate biomarker.

Class C was the second most dominant microstate regardless of the experimental manipulation. There are suggestions from previous research linking it, rather than class D, with self-referential thought (mind-wandering being one instance of it) and the anterior part of the DMN (Michel and Koenig, 2018). Our findings argue against this association, with the C→D transition probability being higher than D→C during mind-wandering as compared to verbalization or visualization (run 2). Furthermore, higher alertness was associated with lower class C *occurrence* and *coverage* during verbalization and lower *occurrence* during visualization, as well as lower probabilities of A→C during verbalization and D→C during verbalization and visualization, whilst greater subjective effort being associated with the higher D→C transition probability during verbalization and visualization. The latter association in particularly seems at odds with the proposed class C association with the DMN-based self-referential processing, as one would expect the opposite transition direction away from the DMN/mind-wandering, rather than towards it, when more subjective effort is employed to stay on task. However, as mentioned earlier, a caveat is that the maps of classes C and D overlap considerably in the previous studies, and were strongly correlated ($r = 0.87$) in this study.

Our findings appear to support Britz and colleagues' (2010) suggestion that class C is related to the function of the saliency network (anterior cingulate and insula), implicated in cognitive control/error detection, which would explain higher probability of transitioning to class C from class D during verbalization and visualization in our experiment, if indeed class D is associated with spontaneous mind-wandering. It does not, however, explain the predominance of class C alongside class D over classes A and B during mind-wandering, as there is no need for error detection in the absence of a task. Likewise, the predominance of class C during wakefulness as well as the lightest and deepest sleep stages in Brodbeck et al.'s study (2012) is also at odds with the idea that class C reflects error detection function. Further research is needed to elucidate the functional significance of class C microstate.

To this end, it might be necessary to further split class C into the classes C and F topographies as described by Custo et al. (2017). Although topographically similar to class C, class F appears to be associated with the activity of the posterior nodes of the DMN. When four canonical classes are used, the anterior and posterior gradient of activation is subsumed under the same class C, which might then reflect the activity of different networks: an anterior saliency network (anterior cingulate and insula), activated during verbalization and visualization in association with error detection/reorienting back to task vs. a posterior network (posterior cingulate/precuneus), known to be associated with mind-wandering/self-referencing and 'captured' by the mind-wandering condition in our experiment.

Contrary to the predictions based on previous research with class A being associated with visual and class B with verbal information processing modalities (Milz et al., 2016), there were no differences between classes A and B on any of the parameters in the verbalization and visualization conditions in the present study. Both classes had the shortest *duration* and lowest *coverage* and *occurrence* compared with classes C and D in all conditions. Both classes had greater *occurrence* during mind-wandering than visualization (run 1).

There are, however, a few indicators for the association of class B with the visual processing modality. Specifically, for the averaged parameters, class B had greater *coverage* during visualization than either mind-wandering or verbalization. Furthermore, transitions to class B (i.e. A→B, C→B or D→B) had higher probability during visualization than either mind-wandering or verbalization (run 1), whereas the B→D transition had a higher probability during mind-wandering and visualization than verbalization (run 2). Additionally, less effort during visualization was associated with longer class B *duration* (run 1). The association of class B with visual imagery is suggested by Brodbeck et al. (2012) findings of class B predominance during stage 2 of non-REM sleep, known to be associated with significant increase of BOLD fluctuations in the visual cortex (Horovitz et al., 2008). It is also in line with its correlation with the positive BOLD signal in the occipital cortex (Britz et al., 2010; Custo et al., 2017), as well as Seitzman et al. (2017) findings of increased class B *coverage* and *occurrence* during eyes-open vs. eyes-closed resting state.

The functional significance of class A microstate is most equivocal in the context of the present study. Milz et al. (2016) linked class A to visuo-spatial processing. Source localization of class A microstate generators has placed them in the posterior part of left-hemisphere language processing areas (Milz et al., 2016; Custo et al., 2017), with the negative BOLD correlation of class A in the same areas (Britz et al., 2010). These associations point to class A topography reflecting an inhibition of language processing, but they do not directly implicate visual processing. We find no clear evidence for the association of class A either with reduced verbal or increased visual information processing, as class A parameters remained low in all conditions.

If anything, our findings suggest against class A association with visual processing, as indicated by the C→A transition probability being the lowest amongst all others during visualization across the runs and repetitions, assuming that class C signals the engagement of the saliency/error detection network with subsequent re-engagement of task-related visual networks and concomitant suppression of posterior

language network indexed by class A topography. Notably, the mean C→A transition probability remained consistently one of the lowest for all repetitions of all conditions, particularly during visualization. On the other hand, A→C transition had higher probability than C→A across all conditions. The associations with the subjective ratings do not provide many clues as to the functional significance of class A either. Higher alertness ratings were associated with longer class A *duration* during mind-wandering and verbalization, whereas higher effortfulness was associated with lower *coverage* during visualization, with both findings further arguing against class A association with visual information processing. The transition probabilities to and from class A had the highest number of associations with the subjective ratings, but the pattern of the observed associations is inconsistent (from run to run and repetition to repetition). Taken together, our findings do not provide support for the association of class A with visual information processing.

4.3. Limitations and future directions

Our findings are based on the EEG data using 30 channels, as opposed to Milz et al. (2016) who used 64-channel recordings. However, Koenig et al. (2002) found the same canonical map classes in the 19-channel EEG data, and Khanna et al. (2014) reported reproducible microstate parameters across 30-, 19- and 8-channel recordings.

We have observed a relatively low total explained variance (69%) by the four canonical microstate classes, compared with approx. 80% reported by the previous studies (e.g. Koenig et al., 1999; Kindler et al., 2011). However, the data-driven topographical maps showed a good correspondence to the four canonical classes (Fig. 1). Notably, the four canonical classes were observed in our verbalization and visualization conditions despite them being a cognitive manipulation/task rather than a 'pure' form of resting state eyes-closed EEG which the four canonical states were historically derived from.

Due to a higher correlation of our class A with Milz et al. (2016) class D ($r = 0.94$) than A ($r = 0.80$), some of the momentary maps that have been classified in the present study as Milz et al. class D would have been classified as class A if data-driven population maps were used. Although not presenting a methodological issue for the test-retest reliability analysis, since the microstate labelling method we have used ensured consistency across runs and repetitions, it might have affected the parameter *coverage* in the present study as compared with that of Milz et al. (2016, Supplementary Table 4), with our class A *coverage* being lower (~18% vs. ~22%) and that of D being higher (~35% vs. ~29%).

The possible effect of an overlap between our class A with Milz et al. class D on estimating the parameters of classes A and D might have affected the investigation of functional significance of classes A and D in our study. However, the overall pattern of findings, discussed in detail earlier, strongly implicates class D in mind-wandering, including i) longer class D *duration* in the mind-wandering than verbalization condition; ii) longer class D *duration* than classes B and C in the mind-wandering than in either verbalization or visualization conditions; iii) an association of lower class D *occurrence* with higher subjective alertness ratings in the mind-wandering condition; and iv) an association of higher degree of spontaneous mind-wandering (as subjectively rated) with higher class D *occurrence* in the verbalization and visualization conditions. Particularly the findings iii) and iv) are noteworthy, as the significance of these correlations would have been even stronger should the estimation of class D parameters have been affected by a significant misclassification of class A as class D maps.

If class A parameters were to be significantly impacted by a possible misclassification as class D maps in our study, we would expect class B parameters to be significantly different than class A parameters in the verbalization condition, if class B is indeed associated with verbalization whilst class A is associated with visualization, as was predicted based on the findings of previous research. Instead, both classes A and B had the shorter *duration* as well as lower *coverage* and *occurrence* than classes C and D in all conditions, with no differences between classes A and B on any of the parameters in either verbalization or visualization conditions. Thus, the functional significance of class A remains uncertain, based on the findings of other studies as well as our own. As rightly noted by Milz et al. (2016), it is premature to associate classes A and B with visual vs. verbal information processing, respectively, as both occur in conditions during which either of these processes is engaged. More research is required to establish the functional significance of classes A and B.

We have used 1–20 Hz band-pass filtered EEG data, following the Milz et al. (2016) protocol. However, it has been suggested (Gärtner et al., 2015) that the actual microstates underlying cognitive processes might be more fleeting and short-lived (about 10 ms of average duration, as compared to 88 ms in the present study), even in 1–40 Hz band-pass filtered EEG data (Brodbeck et al., 2012), and that such short microstates might be difficult to recover from 1 to 20 Hz band-pass filtered EEGs. Moreover, the microstate classes are assigned only to the EEG frames located at the peaks of the GFP curve and are interpolated in between. However, all microstate studies must use strategies to deal with the EEG noise that dominates low-power intervals, whether it be interpolation of microstate classes between GFP peaks (Milz et al., 2016), temporal smoothing (Tomescu et al., 2014), or penalizing transitions from one microstate class to another (Pascual-Marqui et al., 1995).

We note that our two pilot studies had different male/female participant ratios than the main study, with pilot 1 sample consisting of 5 females out of 6 participants, pilot 2 consisting of 3 female and 3 male participants, whereas the final analysed sample of the main study had the male/female ratio of 16/4. However, in pilot 2, which compared 2 s and 4 s intervals for the self-timed repetition rates during verbalization and visualization of 'square', there were no differences noted upon debrief in how comfortable the two self-timed repetition intervals were for male and female participants.

The reported relationships of subjective ratings with the microstate parameters and transition probabilities should be treated as exploratory and interpreted with caution, given the number of performed correlations. However, the observed correlations suggest that the systematic study of the correspondences between the experiential (subjective) dynamics and the neural dynamics as indexed by microstate syntax might be a fruitful avenue to pursue in understanding the functional significance of the EEG microstates. Specifically, there were more associations of the subjective ratings of alertness with the transition probabilities than microstate parameters in our study. Further studies need to be looking at more complex syntax using methods proposed previously (e.g. Nehaniv and Antonova, 2017).

Finally, the relatively low variance explained (69%) when labelling our data using the canonical set of four microstates together with a considerable overlap between the topographies of classes A and D suggest that the EEG microstate research should focus its future efforts on identifying data- and/or population-specific microstates class topographies, as well as developing methods for assessing the optimal number of microstate classes for each dataset (e.g. Pascual-Marqui et al., 1995; Murray et al., 2008; Custo et al., 2017). Attention should also be given to the quantification of within-subject vs. between-subject variations when working with the data-driven EEG microstates (e.g. Zanesco et al., 2020), developing rigorous approaches that would allow for subject-level analysis/statistical inferences alongside more traditional group-level ones (e.g. Nehaniv and Antonova, 2017).

4.4. Conclusions

The present study demonstrates that the differences between information processing modalities (verbal and visual) were smaller than intra-subject differences across the repetitions of the same cognitive manipulation for both microstate parameters and transition probabilities. *Duration* was the most reliable parameter, whereas *occurrence* and transition probabilities between microstate classes were least reliable. The functional significance of the microstates remains unclear, but the pattern of findings suggest that class D might be associated with

spontaneous mind-wandering. The alpha power, the test-retest reliability of microstate parameters and the subjective ratings of alertness and effortfulness indicate that the parameters obtained at the beginning of a 20-min recording session are more robust (e.g. true to the task at hand) and reproducible. This calls for shorter EEG recording sessions with regular breaks to ensure reliability of EEG microstates parameters and transition probabilities for the study of mental processes and clinical populations.

CRediT authorship contribution statement

Elena Antonova: Conceptualization, Methodology, Funding acquisition, Supervision, Formal analysis, Writing – original draft, Writing – review & editing. **Martin Holding:** Investigation, Data curation, Formal analysis, Writing – original draft. **Ho Chak Suen:** Formal analysis, Visualization. **Alex Sumich:** Funding acquisition, Resources, Writing – review & editing. **Reinoud Maex:** Software, Formal analysis, Data curation, Visualization, Writing – review & editing. **Chrystopher Nehaniv:** Conceptualization, Methodology, Funding acquisition, Supervision, Formal analysis, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The study was supported by a research grant from the BIAL Foundation (grant number: 183/16) to Dr Elena Antonova and Prof Chrystopher Nehaniv. The funder had no role in study design; in data collection, analysis and interpretation; in the writing of the report; or in the decision to submit the article for publication. We would like to thank the participants for their contribution to this research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ynirp.2022.100089.

References

Aldridge, V.K., Dovey, T.M., Wade, A., 2017. Assessing test-retest reliability of psychological measures: persistent methodological problems. Eur. Psychol. 22, 207–218.

Al Zoubi, O., Mayei, A., Tsuchiyagaito, A., Misaki, M., Zotev, V., Refai, H., et al., 2019. EEG microstates temporal dynamics differentiate individuals with mood and anxiety disorders from healthy subjects. Front. Hum. Neurosci. https://doi.org/10.3389/fnhum.2019.00056.

Britz, J., Van De Ville, D., Michel, C.M., 2010. BOLD correlates of EEG topography reveal rapid resting-state network dynamics. Neuroimage 52, 1162–1170.

Bréchet, L., Brunet, D., Birot, G., Gruetter, R., Michel, C.M., Jorge, J., 2018. Capturing the spatiotemporal dynamics of task-initiated thoughts with combined EEG and fMRI. bioRxiv. https://doi.org/10.1101/346346.

Brodbeck, V., Kuhn, A., von Wegner, F., Morzelewski, A., Tagliazucchi, E., Borisov, S., Michel, C.M., Laufs, H., 2012. EEG microstates of wakefulness and NREM sleep. Neuroimage 62, 2129–2139.

Coolican, H., 2009. In: Research Methods and Statistics in Psychology, fifth ed. Hodder Education Group, London, United Kingdom.

Cooper, N.R., Croft, R.J., Dominey, S.J., Burgess, A.P., Gruzelier, J.H., 2003. Paradox lost? Exploring the role of alpha oscillations during externally vs. internally directed attention and the implications for idling and inhibition hypotheses. Int. J. Psychophysiol. 47, 65–74.

Croft, R.J., Barry, R.J., 2002. Issues relating to the subtraction phase in EOG artefact correction of the EEG. Int. J. Psychophysiol. 44, 187–195.

Custo, A., Van De Ville, D., Wells, W.M., Tomescu, M.I., Brunet, D., Michel, C.M., 2017. Electroencephalographic resting-state networks: source localization of microstates. Brain Connect. 7, 671–682.

Delorme, D., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. J. Neurosci. Methods 134, 9–21.

Dierks, T., Jelic, V., Julin, P., Maurer, K., Wahlund, L.O., Almkvist, O., Strik, W.K., Winblad, B., 1997. EEG-microstates in mild memory impairment and Alzheimer's disease: possible association with disturbed information processing. J. Neural. Transm. 104, 483–495.

Dinov, M., Leech, R., 2017. Modeling uncertainties in EEG microstates: analysis of real and imagined motor movements using probabilistic clustering-driven training of probabilistic neural networks. Front. Hum. Neurosci. 11, 534.

Gärtner, M., Brodbeck, V., Laufs, H., Schneider, G., 2015. A stochastic model for EEG microstate sequence analysis. Neuroimage 104, 199–208.

Horovitz, S.G., Fukunaga, M., de Zwart, J.A., van Gelderen, P., Fulton, S.C., Balkin, T.J., et al., 2008. Low frequency BOLD fluctuations during resting wakefulness and light sleep: a simultaneous EEG–fMRI study. Hum. Brain Mapp. 29, 671–682.

Horovitz, S.G., Braun, A.R., Carr, W.S., Picchioni, D., Balkin, T.J., Fukunaga, M., et al., 2009. Decoupling of the brain's default mode network during deep sleep. Proc. Natl. Acad. Sci. U. S. A. 106, 11376–11381.

Irisawa, S., Isotani, T., Yagyu, T., Morita, S., Nishida, K., Yamada, K., Yoshimura, M., Okugawa, G., Nobuhara, K., Kinoshita, T., 2006. Increased omega complexity and decreased microstate duration in nonmedicated schizophrenic patients. Neuropsychobiology 54, 134–139.

Khanna, A., Pascual-Leone, A., Farzan, F., 2014. Reliability of resting-state microstate features in electroencephalography. PLoS One 9, e114163.

Khanna, A., Pascual-Leone, A., Michel, C.M., Farzan, F., 2015. Microstates in resting-state EEG: current status and future directions. Neurosci. Biobehav. Rev. 49, 105–113.

Kikuchi, M., Koenig, T., Wada, Y., Higashima, M., Koshino, Y., Strik, W., Dierks, T., 2007. Native EEG and treatment effects in neuroleptic-naïve schizophrenic patients: time and frequency domain approaches. Schizophr. Res. 97, 163–172.

Kindler, J., Hubl, D., Strik, W.K., Dierks, T., Koenig, T., 2011. Resting-state EEG in schizophrenia: auditory verbal hallucinations are related to shortening of specific microstates. Clin. Neurophysiol. 122, 1179–1182.

Koenig, T., Lehmann, D., Merlo, M.C., Kochi, K., Hell, D., Koukkou, M., 1999. A deviant EEG brain microstate in acute, neuroleptic-naive schizophrenics at rest. Eur. Arch. Psychiatr. Clin. Neurosci. 249, 205–211.

Koenig, T., Prichep, L., Lehmann, D., Sosa, P.V., Braeker, E., Kleinlogel, H., Isenhart, R., John, E.R., 2002. Millisecond by millisecond, year by year: normative EEG microstates and developmental stages. Neuroimage 16, 41–48.

Koo, T.K., Li, M.Y., 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J. Chiropr. Med. 15, 155–163.

Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159–174.

Liu, W., Liu, X., Dai, R., Tang, X., 2017. Exploring differences between left and right hand motor imagery via spatio-temporal EEG microstate. Comput. Assist. Surg. 22 (1), 258–266.

Lehmann, D., Ozaki, H., Pal, I., 1987. EEG alpha map series: brain micro-states by space-oriented adaptive segmentation. Electroencephalogr. Clin. Neurophysiol. 67, 271–288.

Lehmann, D., Faber, P.L., Galderisi, S., Herrmann, W.M., Kinoshita, T., Koukkou, M., Mucci, A., Pascual-Marqui, R.D., Saito, N., Wackermann, J., Winterer, G., Koenig, T., 2005. EEG microstate duration and syntax in acute, medication-naive, first-episode schizophrenia: a multi-center study. Psychiatr. Res. 138, 141–156.

Mathôt, S., Schreij, D., Theeuwes, J., 2012. OpenSesame: an open-source, graphical experiment builder for the social sciences. Behav. Res. Methods 44 (2), 314–324. https://doi.org/10.3758/s13428-011-0168-7.

Michel, C.M., Koenig, T., 2018. EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: a review. Neuroimage 180, 577–593.

Milz, P., 2015. Keypy: the KEY EEG analysis toolbox. Zenodo. https://doi.org/10.5281/zenodo.29120. https://zenodo.org/badge/17507/keyinst/keypy.svg.

Milz, P., Faber, P.L., Lehmann, D., Koenig, T., Kochi, K., Pascual-Marqui, R.D., 2016. The functional significance of EEG microstates–Associations with modalities of thinking. Neuroimage 125, 643–656.

Milz, P., Pascual-Marqui, R.D., Achermann, P., Kochi, K., Faber, P.L., 2017. The EEG microstate topography is predominantly determined by intracortical sources in the alpha band. Neuroimage 162, 353–361.

Murray, M.M., Brunet, D., Michel, C.M., 2008. Topographic ERP analysis: a step-by-step tutorial review. Brain Topogr. 20, 249–264.

Nehaniv, C.L., Antonova, E., 2017. Simulating and reconstructing neurodynamics with Epsilon-automata applied to electroencephalography (EEG) microstate sequences. In: Proceedings of IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain (IEEE CCMB'17), IEEE Symposium Series on Computational Intelligence. IEEE Press, pp. 1753–1761.

Nishida, K., Morishima, Y., Yoshimura, M., Isotani, T., Irisawa, S., Jann, K., Dierks, T., Strik, W., Kinoshita, T., Koenig, T., 2013. EEG microstates associated with salience and frontoparietal networks in frontotemporal dementia, schizophrenia and Alzheimer's disease. Clin. Neurophysiol. 124, 1106–1114.

O'Gorman, R.L., Poil, S.S., Brandeis, D., Klaver, P., Bollmann, S., Ghisleni, C., Lüchinger, R., Martin, E., Shankaranarayanan, A., Alsop, D.C., Michels, L., 2013. Coupling between resting cerebral perfusion and EEG. Brain Topogr. 26, 442–457.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9, 97–113.

Pascual-Marqui, R.D., Michel, C.M., Lehmann, D., 1995. Segmentation of brain electrical activity into microstates: model estimation and validation. IEEE Trans. Biomed. Eng. 42, 658–665.

Pascual-Marqui, R.D., Biscay, R.J., Bosch-Bayard, J., Lehmann, D., Kochi, K., Kinoshita, T., Yamada, N., Sadato, N., 2014. Assessing direct paths of intracortical causal information flow of oscillatory activity with the isolated effective coherence (iCoh). Front. Hum. Neurosci. 8, 448.

Post, M.W., 2016. What to do with "moderate" reliability and validity coefficients? Arch. Phys. Med. Rehabil. 97, 1051–1052.

Sämann, P.G., Wehrle, R., Hoehn, D., Spoormaker, V.I., Peters, H., Tully, C., Holsboer, F., Czisch, M., 2011. Development of the brain's default mode network from wakefulness to slow wave sleep. Cerebr. Cortex 21, 2082–2093.
Schlegel, F., Lehmann, D., Faber, P.L., Milz, P., Gianotti, L.R., 2012. EEG microstates during resting represent personality differences. Brain Topogr. 25, 20–26.
Seitzman, B.A., Abell, M., Bartley, S.C., Erickson, M.A., Bolbecker, A.R., Hetrick, W.P., 2017. Cognitive manipulation of brain electric microstates. Neuroimage 146, 533–543.
Spring, J.N., Bourdillon, N., Barral, J., 2018. Resting EEG microstates and autonomic heart rate variability do not return to baseline one hour after a submaximal exercise. Front. Neurosci. 12, 460.
Sten, S., Lundengård, K., Witt, S.T., Cedersund, G., Elinder, F., Engström, M., 2017. Neural inhibition can explain negative BOLD responses: a mechanistic modelling and fMRI study. Neuroimage 158, 219–231.
Stevens, A., Kircher, T., 1998. Cognitive decline unlike normal aging is associated with alterations of EEG temporo-spatial characteristics. Eur. Arch. Psychiatr. Clin. Neurosci. 248, 259–266.
Streiner, D.L., Norman, G.R., 2008. In: Health Measurement Scales: A Practical Guide to Their Development and Use, fourth ed. Oxford University Press, New York.
Strik, W.K., Chiaramonti, R., Muscas, G.C., Paganini, M., Mueller, T.J., Fallgatter, A.J., Versari, A., Zappoli, R., 1997. Decreased EEG microstate duration and anteriorisation of the brain electrical fields in mild and moderate dementia of the Alzheimer type. Psychiatr. Res. 75, 183–191.
Tadel, F., Baillet, S., Mosher, J.C., Pantazis, D., Leahy, R.M., 2011. Brainstorm: a user-friendly application for MEG/EEG analysis. Comput. Intell. Neurosci., 879716
Tomescu, M.I., Rihs, T.A., Becker, R., Britz, J., Custo, A., Grouiller, F., Schneider, M., Debbané, M., Eliez, S., Michel, C.M., 2014. Deviant dynamics of EEG resting state pattern in 22q11.2 deletion syndrome adolescents: a vulnerability marker of schizophrenia? Schizophr. Res. 157, 175–181.
Tomescu, M.I., Rihs, T.A., Roinishvili, M., Karahanoglu, F.I., Schneider, M., Menghetti, S., Van De Ville, D., Brand, A., Chkonia, E., Eliez, S., Herzog, M.H., Michel, C.M., Cappe, C., 2015. Schizophrenia patients and 22q11.2 deletion syndrome adolescents at risk express the same deviant patterns of resting state EEG microstates: a candidate endophenotype of schizophrenia. Schizophr. Res. Cogn. 2, 159–165.
Tomescu, M.I., Rihs, T.A., Rochas, V., Hardmeier, M., Britz, J., Allali, G., Fuhr, P., Eliez, S., Michel, C.M., 2018. From swing to cane: sex differences of EEG resting-state temporal patterns during maturation and aging. Dev. Cogn. Neurosci. 31, 58–66.
von Wegner, F., Tagliazucchi, E., Laufs, H., 2017. Information-theoretical analysis of resting state EEG microstate sequences - non-Markovianity, non-stationarity and periodicities. Neuroimage 158, 99–111.
Whitfield-Gabrieli, S., Ford, J.M., 2012. Default mode network activity and connectivity in psychopathology. Annu. Rev. Clin. Psychol. 8, 49–76.
Zanesco, A.P., King, B.G., Skwara, A.C., Saron, C.D., 2020. Within and between-person correlates of the temporal dynamics of resting EEG microstates. Neuroimage 211. https://doi.org/10.1016/j.neuroimage.2020.116631.