OXFORD

Stimulus-Induced Narrowband Gamma Oscillations are Test–Retest Reliable in Human EEG

Wupadrasta Santosh Kumar[1], Keerthana Manikandan [1], Dinavahi V.P.S. Murty[1], Ranjini Garani Ramesh[1], Simran Purokayastha[1],

Mahendra Javali[2], Naren Prahalada Rao[3] and Supratim Ray[1,*]

[1]Centre for Neuroscience, Indian Institute of Science, Bengaluru 560012, Karnataka, India
[2]MS Ramaiah Medical College & Memorial Hospital, Bengaluru 560012, Karnataka, India
[3]National Institute of Mental Health and Neurosciences, Bengaluru 560012, Karnataka, India
*Address correspondence to Supratim Ray, Centre for Neuroscience, Indian Institute of Science, Bengaluru 560012, Karnataka, India. Email: sray@iisc.ac.in

Abstract

Visual stimulus-induced gamma oscillations in electroencephalogram (EEG) recordings have been recently shown to be compromised in subjects with preclinical Alzheimer's Disease (AD), suggesting that gamma could be an inexpensive biomarker for AD diagnosis provided its characteristics remain consistent across multiple recordings. Previous magnetoencephalography studies in young subjects have reported consistent gamma power over recordings separated by a few weeks to months. Here, we assessed the consistency of stimulus-induced slow (20–35 Hz) and fast gamma (36–66 Hz) oscillations in subjects ($n = 40$) (age: 50–88 years) in EEG recordings separated by a year, and tested the consistency in the magnitude of gamma power, its temporal evolution and spectral profile. Gamma had distinct spectral/temporal characteristics across subjects, which remained consistent across recordings (average intraclass correlation of ∼0.7). Alpha (8–12 Hz) and steady-state-visually evoked-potentials were also reliable. We further tested how EEG features can be used to identify 2 recordings as belonging to the same versus different subjects and found high classifier performance (AUC of ∼0.89), with temporal evolution of slow gamma and spectral profile being most informative. These results suggest that EEG gamma oscillations are reliable across sessions separated over long durations and can also be a potential tool for subject identification.

Key words: EEG; gamma oscillations; healthy aging; intersubject variability; test–retest reliability.

Introduction

Gamma rhythm corresponds to narrowband oscillatory activity between 30 Hz and 70 Hz in the brain, which has been shown to be involved in many higher cognitive functions like attention (Fries et al. 2001; Jensen et al. 2007) and memory (Herrmann et al. 2004). Gamma oscillations have also been shown to be affected in cognitive disorders such as schizophrenia, autism (Uhlhaas and Singer 2010, 2012; Kitchigina 2018), bipolar disorder (Garcia-Rill et al. 2019) and Alzheimer's Disease (AD) (Stam et al., 2002; Kitchigina 2018). In addition, gamma oscillations can be induced by presenting certain visual stimuli such as bars and gratings, and their amplitude and center frequency dependent on the stimulus properties such as size, contrast, and spatial frequency (Henrie and Shapley 2005; Gieselmann and Thiele 2008; Ray and Maunsell 2010; Jia et al. 2013; Murty et al. 2018). Recently, we have shown that these stimulus-induced narrowband gamma oscillations recorded using electroencephalogram (EEG) in middle-aged/elderly subjects (>49 years) weaken with age (Murty et al. 2020) and are weaker in subjects with mild cognitive impairment (MCI) and early AD compared with their age and gender matched controls

(Murty et al. 2021). These results suggest that stimulus-induced gamma oscillations recorded using EEG could potentially be an inexpensive and easily accessible biomarker for diagnosis of mental disorders such as AD.

For the induced gamma to serve as a clinical biomarker, it is necessary to confirm its consistency across multiple recording sessions. Previous studies have shown that stimulus-induced visual gamma oscillations are test–retest reliable, but there are several limitations. First, such test–retest reliability of induced visual gamma oscillations is limited to magnetoencephalogram (MEG) recordings (Hoogenboom et al. 2006; Muthukumaraswamy et al. 2010; Tan et al. 2016). The only EEG study to our knowledge that reported induced gamma oscillations used a checkerboard stimulus that mainly induced broadband gamma (Keil et al. 2003); another study that used gratings to induce gamma oscillations reported reliability of evoked, not induced, gamma oscillations (Fründ et al. 2007). Second, although having a test–retest interval of a year allows for checking its feasibility as an annual check-up, most of these studies have tested reliability with test–retest interval spanning from a few weeks to months (but see

McCusker et al. 2021 who had a test–retest interval of 3 years, although they used checkerboard patterns and their paradigm also included a behavioral task). Third, most studies have been conducted on young subjects (20–40 years), but it is also important to study the reliability in older subjects who are prone to cognitive disorders. Further, steady-state-visually evoked-potentials (SSVEPs) at gamma frequencies have recently been shown to have therapeutic effect on rodent models of AD (Iaccarano et al., 2016), and to be useful in brain computer interfacing (BCI) applications (Vialatte et al. 2010). Therefore, it is important to compare the test–retest reliability of SSVEPs with induced gamma oscillations. Finally, most previous studies have focused on power and center frequency of induced gamma oscillations. However, these oscillations also have a characteristic temporal evolution, and in addition, the peak frequencies of these and other oscillations vary across individuals, thus generating a characteristic spectral profile as well. Since these spectral and temporal features could provide additional information beyond simply the power and center frequency, it is important to study test–retest reliability of these features. Such power and center frequency based features have been studied in the context of subject identifiability (Näpflin et al. 2007) and biometric authentication (Thomas and Vinod 2018) in resting state EEG data as well as visual evoked potentials (Sarnthein et al. 2009), but not for stimulus-induced gamma oscillations.

To address these limitations, we tested the reliability of visually induced narrowband gamma oscillations recorded using EEG in healthy middle-aged/elderly (aged >49 years) with an interval of 1 year or more. Since several previous studies have reported reliability of alpha oscillations (Salinsky et al. 1991; Corsi-Cabrera et al., 2007; McCusker et al. 2021), we performed these analyses on the alpha band as well. This dataset is a part of double-blind case–control EEG study ($n = 257$; 106 females), part of which has been used earlier to show that gamma oscillations weaken with healthy aging (Murty et al. 2020) and with MCI/AD (Murty et al. 2021). Fifty-two subjects were retested with the same visual paradigm after a period of approximately 1 year. We employed large visual stationary Cartesian grating stimuli that induced 2 distinct gamma oscillations in both monkeys and humans (Murty et al. 2018), called slow [20–35 Hz] and fast gamma [36–66 Hz]. We estimated reliability of power in alpha and both gamma bands using a widely used measure called Intraclass Correlation Coefficient (ICC; (Shrout and Fleiss 1979)). We examined the test–retest reliability of SSVEPs at 32 Hz as well. We also studied correlations in the temporal and spectral profiles across same versus different individuals. In addition, we studied intersubject variability (subject distinctiveness) through a linear classifier that was trained to discern same subject pairs from the all possible cross-subject pairs and compared the contribution of various features toward classification.

Materials and Methods
Human Subjects
The full dataset was collected from 257 middle-aged/elderly human subjects (females: 106) aged 50–88 years as part of the Tata Longitudinal study of aging (TLSA). For brevity, we refer to the whole population as "elderly," even though there is considerable age variation in the population. The subjects were recruited from urban Bengaluru communities, with careful evaluation by trained psychiatrists, psychologists, and neurologists affiliated with National Institute of Mental Health and Neuroscience (NIMHANS) and M.S. Ramaiah Hospital, Bengaluru. Their cognitive performance was evaluated by a set of tests—Addenbrooke's Cognitive Examination (ACE), Clinical Dementia Rating (CDR), and Hindi Mental State Examination (HMSE). More details have been provided in previous reports (Murty et al. 2020, 2021). Out of the 257 subjects recruited in the first year, we discarded data of 10 subjects due to noise. Out of the remaining 247 subjects, 52 subjects were retested after approximately a year (this test–retest interval was based on constraints not related to EEG recordings, since this was part of a multi-investigator project involving many types of recordings). Among these, 4 subjects were either unhealthy (CDR > 0) or had invalid CDR score, leading to usable data of 48 subjects who were healthy in both sessions. The test–retest sessions of a subject are referred to as baseline and follow-up.

All subjects received monetary compensation for participating in both the sessions. Informed consent was obtained from the subjects prior to the experiment. All procedures were approved by The Institute Human Ethics Committees of Indian Institute of Science, NIMHANS, Bengaluru and M.S. Ramaiah Hospital, Bengaluru.

Experimental Setting and Behavioral Task
Experimental setup and details of recordings have been explained in detail previously (Murty et al. 2020, 2021) and are summarized here. EEG was recorded from 64-channel active electrodes (actiCap) using BrainAmp DC EEG acquisition system (Brain Products GMbH) and were placed according to the international 10–10 system. Raw signals were filtered online between 0.016 Hz (first-order filter) and 1 kHz (fifth-order Butterworth filter), sampled at 2.5 kHz, and digitized at 16-bit resolution (0.1 μV/bit). Samples were decimated to 250 Hz using the Matlab command "resample," which applies an finite impulse response antialiasing lowpass Kaiser filter before down-sampling by 10. Average impedance of the final set of electrodes was (mean ± SEM) 7.89 ± 0.67 kΩ and 7.73 ± 0.54 kΩ for baseline and follow-up, respectively. All the EEG signals recorded were referenced to electrode FCz during acquisition (unipolar reference scheme).

All subjects sat in a dark room facing a gamma-corrected LCD monitor (dimensions: 20.92 × 11.77 inches; resolution: 1280 × 720 pixels; refresh rate: 100 Hz; BenQ XL2411) with their head supported by a chin rest.

It was placed at (mean ± SD) 58 ± 0.7 cm from the subjects (range: 54.9–61.0 cm) based on their comfort and subtended 52° × 30° of visual field for full screen gratings.

In the "gamma" experiment, subjects performed a visual fixation task, wherein 2–3 full screen grating stimuli were presented for 800 ms with an interstimulus interval of 700 ms after a brief fixation period of 1000 ms in each trial using a customized software running on MAC OS. The stimuli were full contrast sinusoidal luminance achromatic gratings with either of the 3 spatial frequencies (1, 2, and 4 cycles per degree (cpd)) and 4 orientations (0°, 45°, 90°, and 135°). Out of the 48 subjects that were retested, 1 subject had noisy follow-up data, leaving with 47 pairs of subjects. With further constraints on the number of good electrodes for analysis (see Artifact Rejection below) over both baseline and follow-up, we finally had 40 pairs of subjects (20 males and 20 females). Subjects with data available on both the sessions ($n = 40$) aged in the range of (mean ± SD) 64.3 ± 7.17 (min: 50, max: 85 years). The test–retest interval was (mean + SD) was 380.4 ± 10.12 days (min: 256, max: 519), as reported in the plots in Figures 1 and 2.

The "SSVEP" experiment was performed on a subset of these subjects. This was always conducted after the gamma experiment, on the same day because the stimulus settings were dependent on the gamma experiment results. Here, subjects viewed a random presentation of grating counter-phasing at 16 cycles per second (cps), along with a static grating, both of spatial frequency, and orientation that produced the maximum gamma in the gamma experiment, but only the data from counter-phased grating was used. These stimuli were presented in a similar stimulus paradigm as the gamma experiment. Out of the 48 subjects that were retested, 3 subjects had unusable baseline or follow-up data and 9 subjects did not have enough number of good electrodes (see Artifact Rejection below), leaving us with 36 pairs of subjects (20 males and 16 females). Of these, subjects with the number of trials less than 15 were further discarded because an appreciable SSVEP peak was often not visible with fewer trials (4 subjects), ending up with 32 pairs (17 males and 15 females).

Eye position was monitored using EyeLink 1000 (SR Research Ltd), sampled at 500 Hz. Any eye-blinks or change in eye position outside a 5° fixation window during −0.5 s to 0.75 s from stimulus onset were noted as fixation breaks, which were removed offline. This led to a rejection of 14.6 ± 2.8% (mean ± SEM) and 17.8 ± 2.9% (mean ± SEM) repeats for the baseline and the follow-up groups for the gamma experiment, and 14.9 ± 2.4% and 16.5 ± 2.7% for the SSVEP experiment. We have previously shown that eye-movements and microsaccades (Hassler et al. 2011) have a negligible effect on gamma power in this dataset (Murty et al. 2020).

Artifact Rejection

We implemented a fully automated artifact rejection framework for further details, see (Murty et al. 2020, 2021), in which outliers were detected as repeats with deviation from the mean signal in either time or frequency domains by more than 6 standard deviations, and subsequently electrodes with too many (30%) outliers were discarded. Subsequently, stimulus repeats that were deemed bad in the visual electrodes (see the section on EEG data analysis below) or in more than 10% of the other electrodes were considered bad, eventually yielding a set of good electrodes and common bad repeats for each subject. Overall, this led us to reject (mean ± SEM) 14.7 ± 0.8% and 14.6 ± 0.8% of repeats for the baseline and follow-up groups for the gamma experiment, and 6.9 ± 0.6% and 7.3 ± 0.7% for the SSVEP experiment. Finally, we computed slopes for the baseline power spectrum between 56 Hz and 84 Hz range for each unipolar electrode and rejected electrodes whose slopes were less than 0. We further discarded protocols which did not have any of the bipolar electrode pair (see EEG data analysis) in the visual anterolateral electrode groups (both left and right anterolateral group) after electrode rejection. Subjects with no good protocols (in either baseline or follow-up group or both) were rejected from analysis, yielding 40 subjects (20 females) for the gamma experiment and 32 subjects for the SSVEP experiment.

In addition to this artifact detection pipeline that was used in previous studies, 2 additional conditions were included. First, for each subject, we computed the union of bad electrodes in baseline and follow-up sessions and rejected this union from both sessions, such that analyses for baseline and follow-up were over the same set of electrodes. Second, we noticed the presence of a large discontinuity in the signal in a small fraction (∼3%) of stimulus repeats and added a pipeline to remove those. Specifically, stimulus repeats with discontinuities (in any of the electrodes) were detected by thresholding the derivative of signal (>120 μV/sec) in time within [−0.5 1.5] s with respect to stimulus onset. This led to the rejection of (mean ± SEM) 5.07 ± 0.89 (maximum: 21) repeats in baseline and 4.55 ± 1.04 (maximum: 23) repeats in follow-up subjects. Electrodes with more than 20 such discontinuities were considered noisy (maximum: 5 electrodes). These additional conditions were added mainly to improve the time–frequency plots (Figs 1 and 2) that were sensitive to these large discontinuities; the main results (Figs 5 and 8) remained similar even without these additional conditions. Overall, after discarding all bad repeats, 291.6 ± 12.8 and 298.9 ± 14.12 stimulus repeats were available for the gamma experiment, and 30.1 ± 1.1 and 29.5 ± 1.1 for the SSVEP experiment, for the baseline and the follow-up groups, respectively.

The exclusion criteria are summarized in Table 1 below. Note that the reduction in the size of the SSVEP dataset is simply because it was always done after
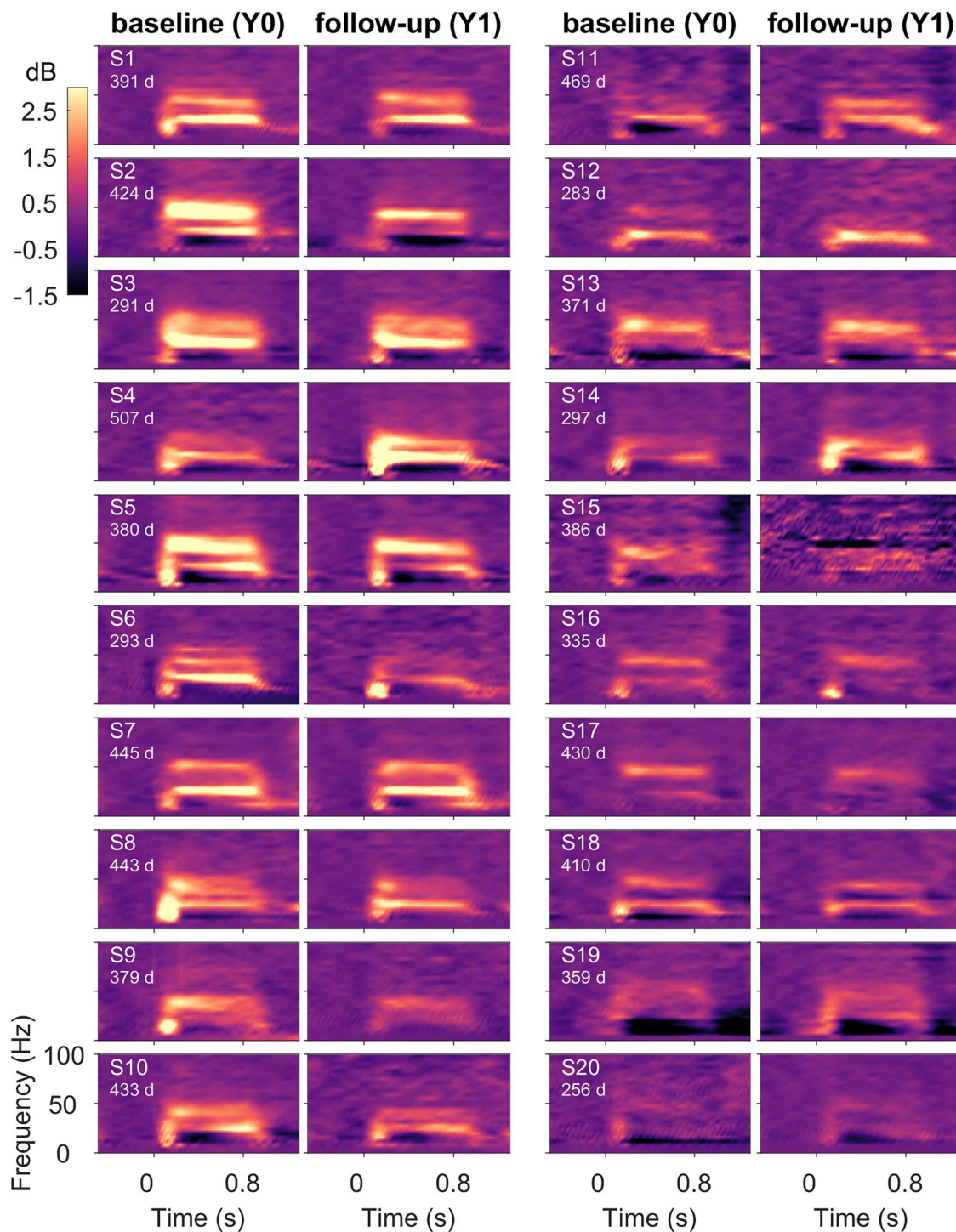
Figure 1. Gamma oscillations remain consistent across baseline and follow-up sessions. Change in time–frequency power spectra for the 2 recording sessions, baseline (left), and follow-up (right) of 20 female subjects to static gratings ("gamma" experiment). Subjects are ordered vertically based on the decreasing average slow and fast gamma power in visual electrodes, starting from the top left. The color bar on the top left denotes the log power ratio in dB units. The numbers indicate test–retest interval in days.

the Gamma experiment and for a shorter duration, since the main aim of this project was to investigate stimulus-induced gamma oscillations. However, even for the Gamma study, data from 8/48 (~17%) of subjects were discarded. This relatively high number of rejections is because of the very stringent criteria for subject selection. For example, the visual electrodes that were used for analysis (see "EEG data analysis" section below) were divided into 2 anterolateral groups and one posteromedial group. We rejected any subject for which even one of the 2 anterolateral groups did not have any good electrodes. This was done so that we could test for potential asymmetry in gamma distribution on the scalp (as shown in Supplementary Figs S1 and S2).
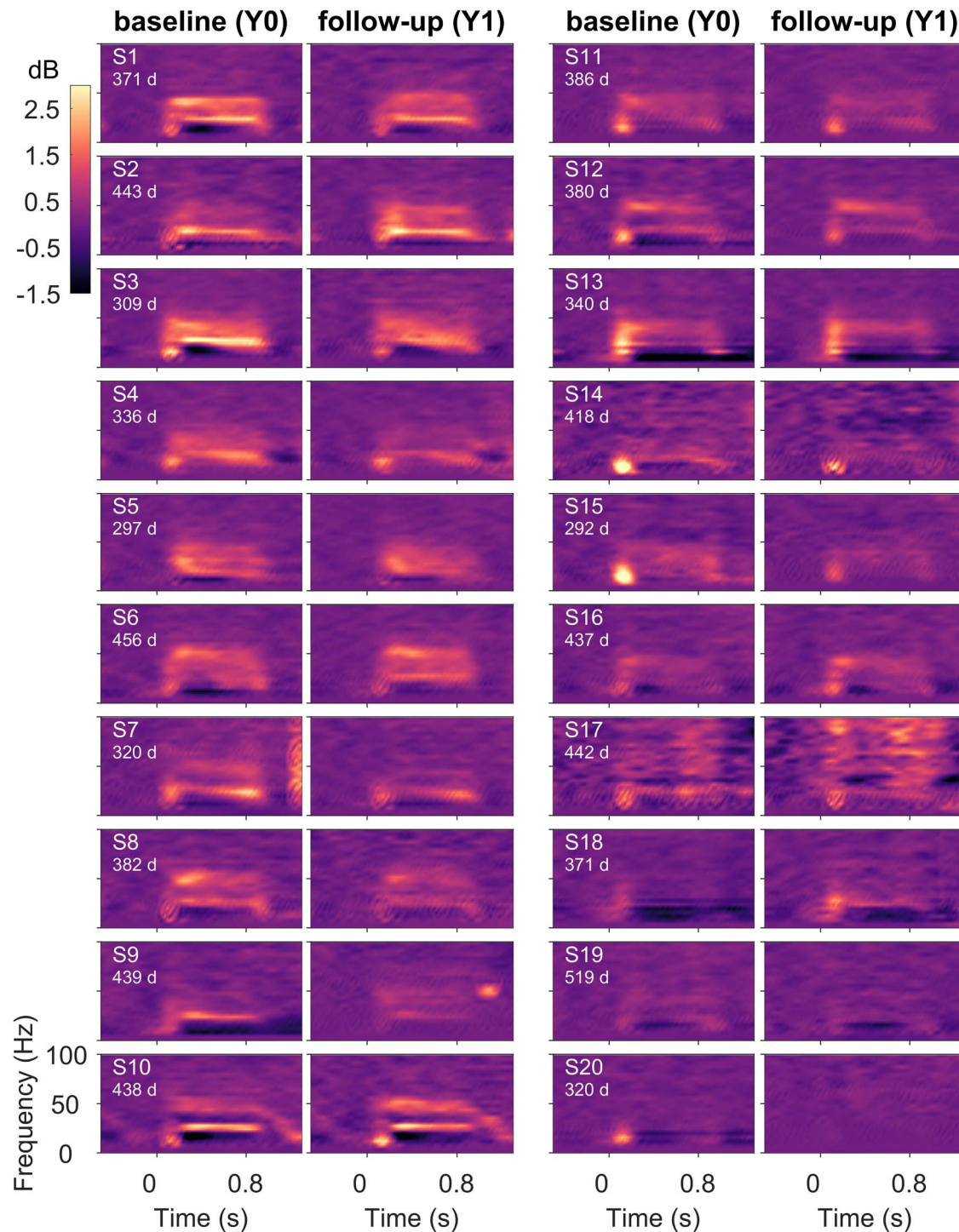
Figure 2. Same as Figure 1 for 20 male subjects.

Relaxing this requirement such as usable electrodes in either one (instead of both) of the 2 anterolateral groups lead to subject selection would have yielded 4 additional subjects (and a rejection ratio of ~8%).

EEG Data Analysis

All analyses were performed using bipolar reference scheme. Every electrode was re-referenced offline to its neighbor, yielding 112 bipolar pairs of electrodes from the 64 unipolar electrodes (Murty et al. 2020, 2021). The

bipolar pairs among the visual electrodes considered for analysis (except for scalp maps) were: PO3-P1, PO3-P3, POz-PO3 (left anterolateral group); PO4-P2, PO4-P4, POz-PO4 (right anterolateral group) and Oz-POz, Oz-O1, Oz-O2 (posteromedial group).

All the data analyses were done using custom codes written in MATLAB (MathWorks. Inc; RRID:SCR_001622). Power spectrum and time–frequency spectrograms were obtained using the Chronux Toolbox ((Bokil et al. 2010), RRID:SCR_005547). Both were calculated for individual

Table 1. The summary of the exclusion criteria for subject rejection across the two experiments, responsible for different subject counts in analysis

S.no	Exclusion criteria	"gamma"	"SSVEP"
1	Not healthy (CDR > 0)	1	1
2	Invalid CDR score	3	3
3	Incomplete experiment or could not be extracted	0	3
4	Unusable baseline or follow-up data	7	9
5	Outlier subjects (manual)	1 [noisy Time Frequency spectra]	4 [weak SSVEP peak] ($n < 15$)
	Overall usable subjects	40	32

trials and then averaged across trials. For analysis, similar to our previous study (Murty et al. 2020), we chose [−500 0] ms as the baseline period and [250 750] ms as the stimulus period, yielding a resolution of 2 Hz. This stimulus period was so chosen to avoid stimulus-onset transients. Time–frequency spectrogram was computed with 4 Hz frequency resolution (moving window of 250 ms) and a step size of 25 ms.

Change in power between stimulus period and baseline period for a frequency band was computed using the following formula:

$$\Delta Power = 10 \left(\log_{10} \frac{\sum_f ST(f)}{\sum_f BL(f)} \right)$$

where ST is the stimulus power spectra and BL is the baseline power spectra, both averaged within relevant frequency bins (f), across all the analyzable repeats. Band powers are specifically computed in 3 frequency bands, namely slow gamma (20–35 Hz), fast gamma (36–66 Hz), and alpha (8–12 Hz). The gamma frequency band limits were set based on the localization of 2 gamma peaks in the change in PSD profile, and are the same as used in our previous studies (Murty et al. 2018, 2020, 2021). These were averaged across all the 3 electrode groups. Scalp maps were generated using the *topoplot* function of EEGLAB toolbox ((Delorme and Makeig 2004), RRID:SCR_007292) with custom bipolar montage of 112 channels.

Test–retest reliability was assessed by Intraclass Correlation Coefficient (ICC) measure. Suppose the dataset is arranged in a matrix with each column representing a session and each row a different subject. ICC incorporates both intersubject and intrasubject variability in terms of mean square (variance) for rows (MSR) [variance of sessions' means across subjects] and mean square (variance) within raters (MSW) [mean of subjects' variance across sessions], respectively. A specific type of ICC, called 1–1 ICC, which is based on one-way analysis of variance model, is defined as follows (see (Shrout and Fleiss 1979) for details):

$$ICC = \frac{MSR - MSW}{MSR + (N_{sessions} - 1) * MSW}$$

In our case, since $N_{sessions} = 2$, this reduces to ICC = (MSR − MSW)/(MSR + MSW). This measure increases with decreasing within-subject variability (MSW) and with increasing between-subject variability (MSR), with values approaching 1 with perfect reliability (MSW = 0). This measure was used for the reliability of spectral power in sensor and source domains in MEG (Martín-Buro et al. 2016). Also, the null-hypothesis H_0: ICC = 0 was tested for significance by F-statistical test, F-value = MSR/MSW, with ($N_{subjects} - 1$) and $N_{subjects}(N_{sessions} - 1)$ degrees of freedom (Shrout and Fleiss 1979), reported with 95% confidence interval, F-value, and P-value. Larger the variation of session means across subjects, relative to the variation within sessions, larger is the F-value. Given the F-value, degrees of freedom of MSR and MSW, P-value is computed based on the F cumulative distribution function. Two related types of ICC, called Consistency and Agreement ICC (ICC(C,1) and ICC(A,1)) were also computed for comparing our results with other studies. These ICC definitions involve mean square for error which encompasses variance across both subjects and sessions (for details, see (McGraw and Wong 1996)). We computed the ICC using MatlabCentral file-exchange codes provided by Salarian (2016), which implemented the statistical testing described by (McGraw and Wong 1996).

Estimating Subject Distinctiveness Using Linear Discriminant Analysis

A measure of subject distinctiveness reflects the separability between the self-pairs and cross-pairs—which is approximated here, in terms of classifier performance (Linear Discriminant classifier) in discriminating the 2 categories of pairs. Input to the classifier is a pair of baseline and follow-up session recordings from 2 subjects. For each pair, we defined 7 comparative measures (stimulus-induced features), namely, absolute difference in relative band powers within slow gamma ($p_{s\gamma}$) [20–35 Hz], fast gamma ($p_{f\gamma}$) [36–66 Hz], and alpha (p_α) [8–12 Hz] (3 features), baseline corrected band power temporal correlations (Pearson's correlation) in the slow gamma ($tc_{s\gamma}$), fast gamma ($tc_{f\gamma}$), and alpha (tc_α) between [0 0.8] s (3 features; correlation between the traces shown in Fig. 3), and spectral correlation (*sc*) in [0-100 Hz] (excluding 50 Hz and 100 Hz peaks that represented line noise and monitor refresh rate; 1 feature; correlation between the traces shown in Fig. 4). To study the importance of the features in the baseline period, we also considered baseline alpha difference power (p_{bl_α}) and raw PSD correlation (*bl sc*) features (correlation between the baseline PSDs shown in Supplementary Fig. S3). Three binary classifiers (using only stimulus-induced features, only baseline features, and both feature sets combined) were trained to classify a pair into either self-pair or cross-pair and was tested according to 5-fold stratified cross-validation.

Because the cross-pairs ($40 \times 39 = 1560$) far outnumbered the self-pairs (40), classifier performance could
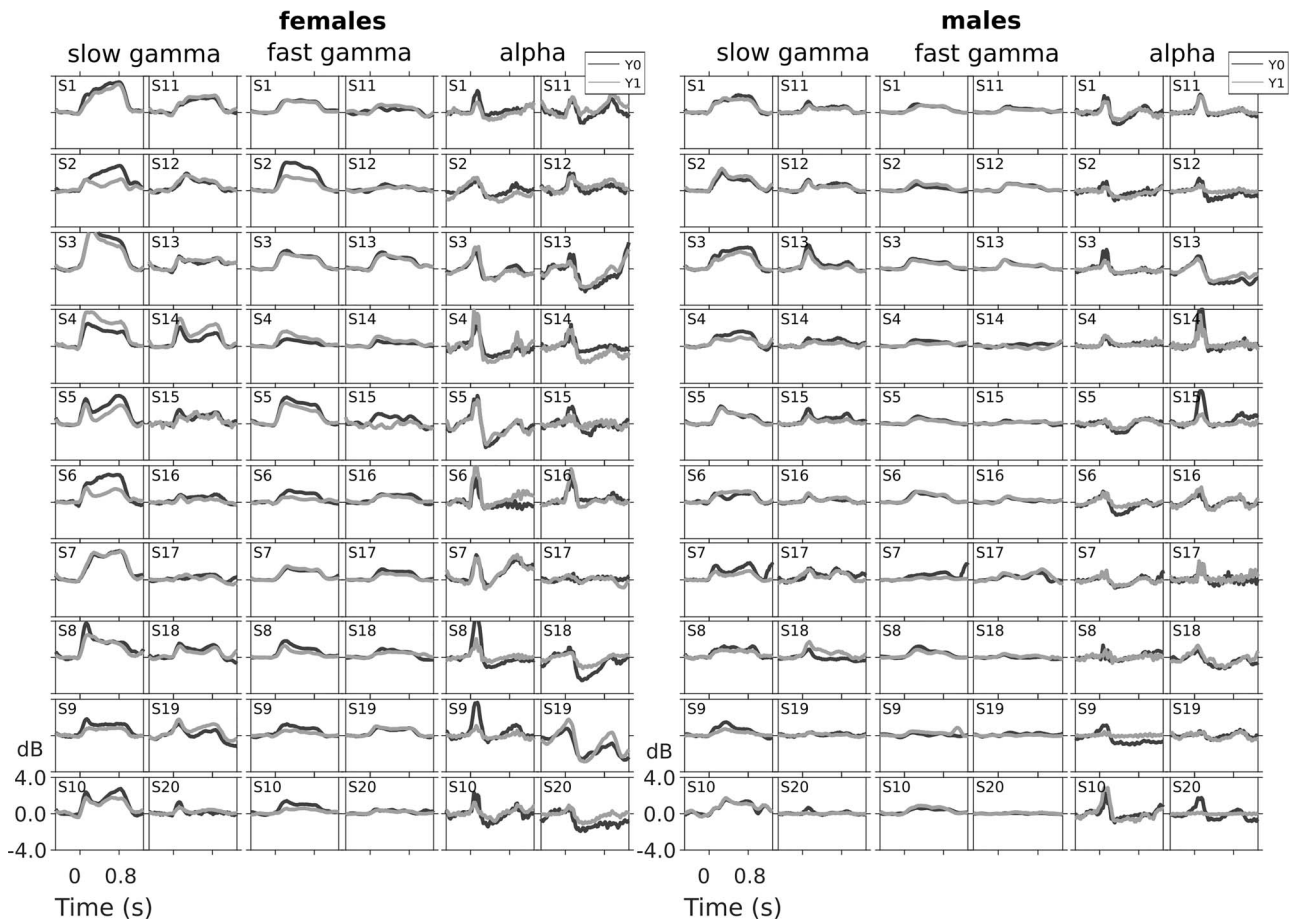
Figure 3. Temporal evolution of band powers in slow gamma, fast gamma, and alpha range across baseline and follow-up remain consistent in males and females. Band power as a function of time, computed from averaging baseline corrected time–frequency spectra across relevant frequency bins for slow gamma, fast gamma, and alpha bands. Traces are plotted in black for baseline (Y0) and gray for follow-up (Y1) for 20 female and 20 male subjects. Same subject order as in previous plots.
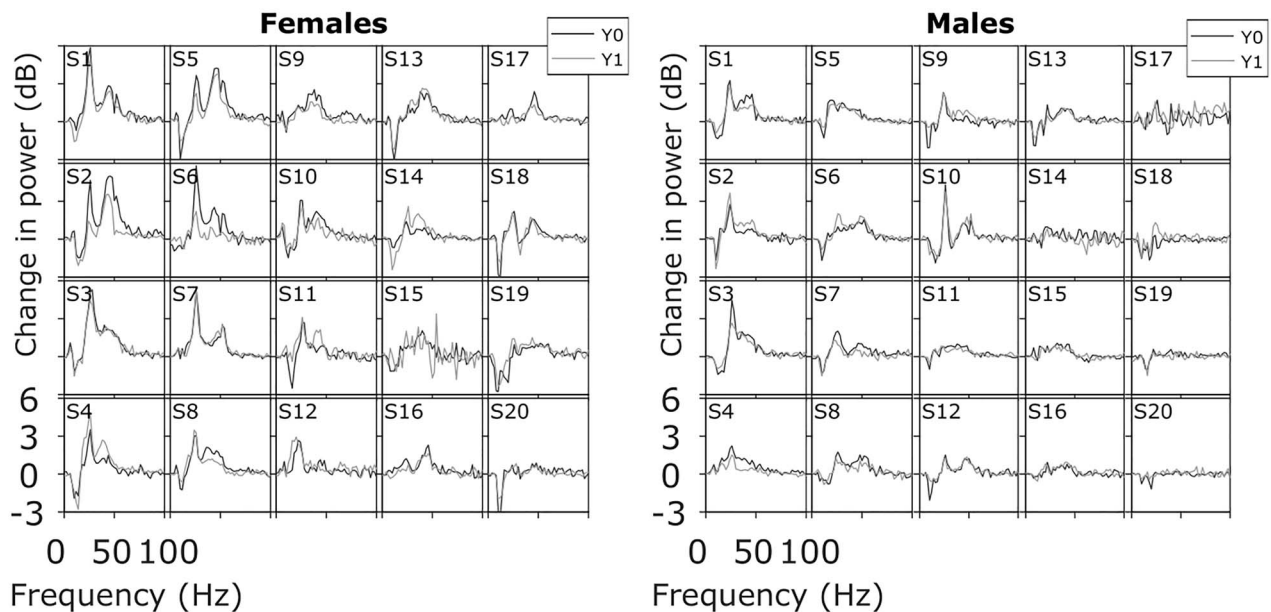

Figure 4. Power spectral densities (PSDs) are consistent across baseline and follow-up in males and females. The change in PSD during stimulus period relative to baseline PSD for baseline (Y0; black) and follow-up (Y1; gray) for 20 female and 20 male subjects.

not be assessed using accuracy, since even a null classifier that categorizes every observation as a cross-pair would have an accuracy of 97.5%. Instead, we assessed the classifier performance through the area under the receiver operating characteristic (ROC) curve (AUC), where the ROC curve represents the true positive rate against false positive rate. AUC calculation does not need the class label output of classifier, as it considers various thresholds for classification. Only the weighted projection along with the true class labels are needed for true positive rate and false positive rate. AUC values vary between 0.5 (chance level classification) and 1 (perfect classification). Feature importance was determined by the weight of the feature computed using linear discriminant analysis (LDA) classifier on the z-scored feature values, as well as using single feature AUC method wherein each feature was individually considered to classify self versus cross-pair. We used MATLAB statistics and machine learning toolbox function *perfcurve.m* for computing AUC.

Statistical Analysis

All the statistical analysis and results obtained, are using the power spectrum or correlation measures. Appropriate tests (t-statistic; F-statistic; Wilcoxon rank sum) were used to interpret our findings.

Data and Code Availability

All spectral analyses were performed using Chronux toolbox (version 2.10), available at http://chronux.org. ICC analysis functions were taken from MatlabCentral file-exchange, provided by Salarian (2016). Figure data and codes used here are available at https://github.com/supratimray/TLSAEEGProjectPrograms under the consistencyProject folder. Raw data will be made available to readers upon reasonable request and made publicly available at a later time.

Results
Time Frequency Spectra Remain Reliable over 1-Year Interval

We first examined the change in time–frequency power spectra relative to baseline power (−0.5 to 0 s) of EEG signals for the "gamma" experiment, averaged across the 3 bipolar electrode groups, for the baseline and follow-up sessions (see Methods). Figures 1 and 2 show the results for female (n = 20) and male subjects (n = 20), sorted based on decreasing total gamma power. Both females and males show visually reliable time–frequency spectra within a subject across baseline (left panel) and follow-up (right panel). Specifically, the powers in the 3 frequency bands, slow gamma, fast gamma, and alpha were consistent in time–frequency spectra, as were the temporal evolution of these rhythms. Similarly, the topographic scalp maps of average gamma band power (mean of slow and fast gamma) appeared consistent between baseline and

follow-up sessions in females (Supplementary Fig. S1) and males (Supplementary Fig. S2).

Band Powers, Spectral and Temporal Profiles All Remain Highly Reliable

Figure 3 shows the temporal evolution of band powers in alpha, slow gamma, fast gamma range, computed by averaging time frequency spectra across relevant frequency bins, separately for males and females. Remarkably, these profiles were distinct across subjects but appeared highly correlated across the 2 sessions within subjects. Figure 4 shows the change in power during the stimulus period (0.25–0.75 s, where 0 is the stimulus onset) relative to baseline for the 2 sessions. These traces show a prominent suppression at alpha range, and elevation of power at slow and fast gamma bands, and together represent a "spectral profile" for each subject. These spectral profiles were also highly similar for each subject and distinct across subjects. Consistency was found in the spectral profile of baseline period absolute power as well, although for several subjects there was substantial difference in the noise floor (Supplementary Fig. S3). For these subjects, the stimulus-induced PSDs were also different in a similar way (data not shown), such that the change in PSD plots shown in Figure 3 remained highly overlapping. Further, in spite of the difference in noise floor, the spectral shape during baseline periods was informative about subject identity (see Fig. 8 for details).

Figure 5a–c shows a scatter plot of power in slow gamma, fast gamma, and alpha bands in baseline versus follow-up sessions. Supplementary Figure S4 shows similar scatter plot of average resting state alpha power (baseline period). SSVEP power (increase in power at 32 Hz from baseline) computed from the SSVEP experiment (see Methods) is also shown (Fig. 5d). In all cases, the powers were highly correlated (significance results are indicated in the plots). Overall, we did not find age dependence or the dominance of any age group in the observed correlation pattern (not shown).

We further assessed the reliability using a metric called ICC (Shrout and Fleiss 1979), which compares the ratio of variances within and across subjects. Average band powers were highly reliable for all frequencies ranges: slow gamma ICC: 0.74 (95% CI = 0.56–0.85; $F = 6.77$, $P = 9.68\mathrm{e}{-9}$), fast gamma ICC: 0.69 (95% CI = 0.50–0.82; $F = 5.64$, $P = 1.41\mathrm{e}{-7}$), alpha ICC: 0.65 (95% CI = 0.43–0.79; $F = 4.68$, $P = 1.89\mathrm{e}{-6}$), and SSVEP ICC: 0.58 (95% CI = 0.30–0.77; $F = 3.75$, $P = 1.76\mathrm{e}{-4}$). In addition, average resting state alpha power was also reliable; ICC: 0.78 (95% CI = 0.62–0.88; $F = 8.12$, $P = 5.94\mathrm{e}{-10}$). Previous studies have used 2 other variants of ICC, called consistency ICC(C,1) and agreement ICC(A,1), see Methods and (McGraw and Wong 1996) for details. These measures were also comparable to the classical ICC in our dataset. For slow gamma, fast gamma, alpha, and SSVEP, ICC(C,1) values were 0.74, 0.72, 0.67, and 0.59, respectively, whereas ICC(A,1) values were 0.74, 0.70, 0.65, and 0.58.
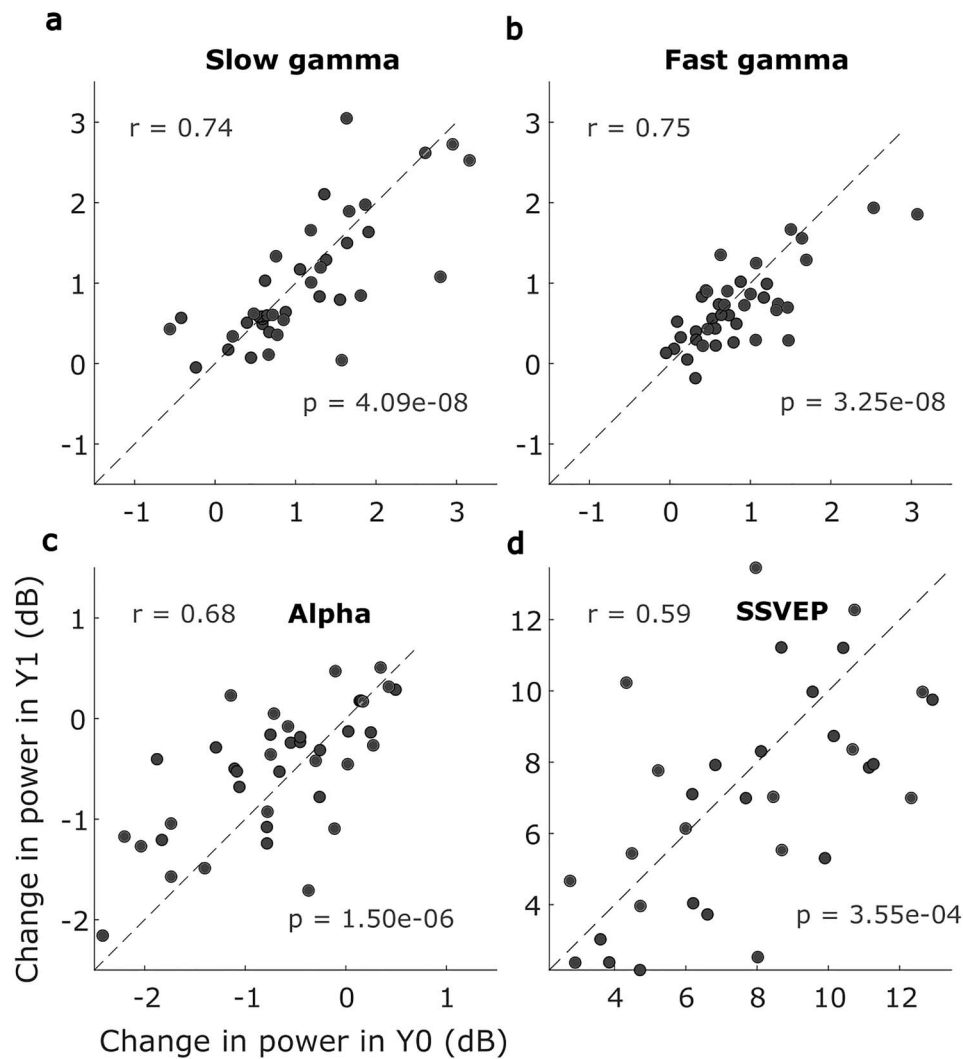
Figure 5. Change in slow gamma, fast gamma, alpha power and SSVEP at 32 Hz power are correlated across baseline and follow-up. (A)–(C) Scatter plots of the average change in power within slow gamma, fast gamma, and alpha bands for baseline and follow-up sessions. (D) The average change in power at SSVEP frequency at 32 Hz. Pearson's correlation coefficient and P-value are mentioned on the top left and bottom right in each panel. Note the difference in axis limits across the plots.

We also assessed the internal consistency of these measures within baseline and follow-up sessions by splitting the trials of a single session into 2 halves. We found band powers to be highly reliable in the 2 halves in both baseline and follow-up sessions for all frequency ranges. Baseline sessions: slow gamma ICC: 0.90 (95% CI = 0.83–0.95; $F = 19.83$, $P = 1.11e-16$), fast gamma ICC: 0.94 (95% CI = 0.89–0.97; $F = 31.55$), alpha ICC: 0.80 (95% CI = 0.66–0.89; $F = 9.30$, $P = 6.71e-11$). Follow-up sessions: slow gamma ICC: 0.94 (95% CI = 0.88–0.97; $F = 30.77$), fast gamma ICC: 0.90 (95% CI = 0.82–0.94; $F = 18.65$, $P = 4.44e-16$), alpha ICC: 0.86 (95% CI = 0.75–0.92; $F = 12.97$, $P = 2.53e-13$).

To quantitatively assess the reliability in the temporal evolution and spectral profile, we computed the Pearson's correlation between baseline and follow-up traces for each subject ("self-pair correlation") and compared with the average correlation between the follow-up traces of all other subjects ("cross-pair

correlation"). This is similar to the analysis done by Vázquez-Marrufo et al. (2017) in which correlation of alpha ERD topography profiles within a subject was compared with the correlation across subjects. For the temporal correlation of slow gamma, self-pair correlation was 0.87 ± 0.09 (median ± median absolute deviation (MAD)), significantly greater than cross-pair correlation (0.28 ± 0.09; left-sided Wilcoxon rank sum test, $P = 1.59e-13$, $n = 40/1560$; Fig. 6a). Similarly, for the temporal correlation of fast gamma, self-pair correlation was 0.88 ± 0.09 (median ± MAD), significantly greater than cross-pair correlation (0.73 ± 0.07; left-sided Wilcoxon rank sum test, $P = 8.74e-7$, $n = 40/1560$; Fig. 6b). Likewise, alpha temporal correlation within self-pair was 0.94 ± 0.03 (median ± MAD), significantly greater than cross-pair correlation (0.88 ± 0.03; left-sided Wilcoxon rank sum test, $P = 2.1e-7$, $n = 40/1560$; Fig. 6c). For spectral correlation, self-pair correlation was 0.85 ± 0.08 (median ± MAD), significantly greater than
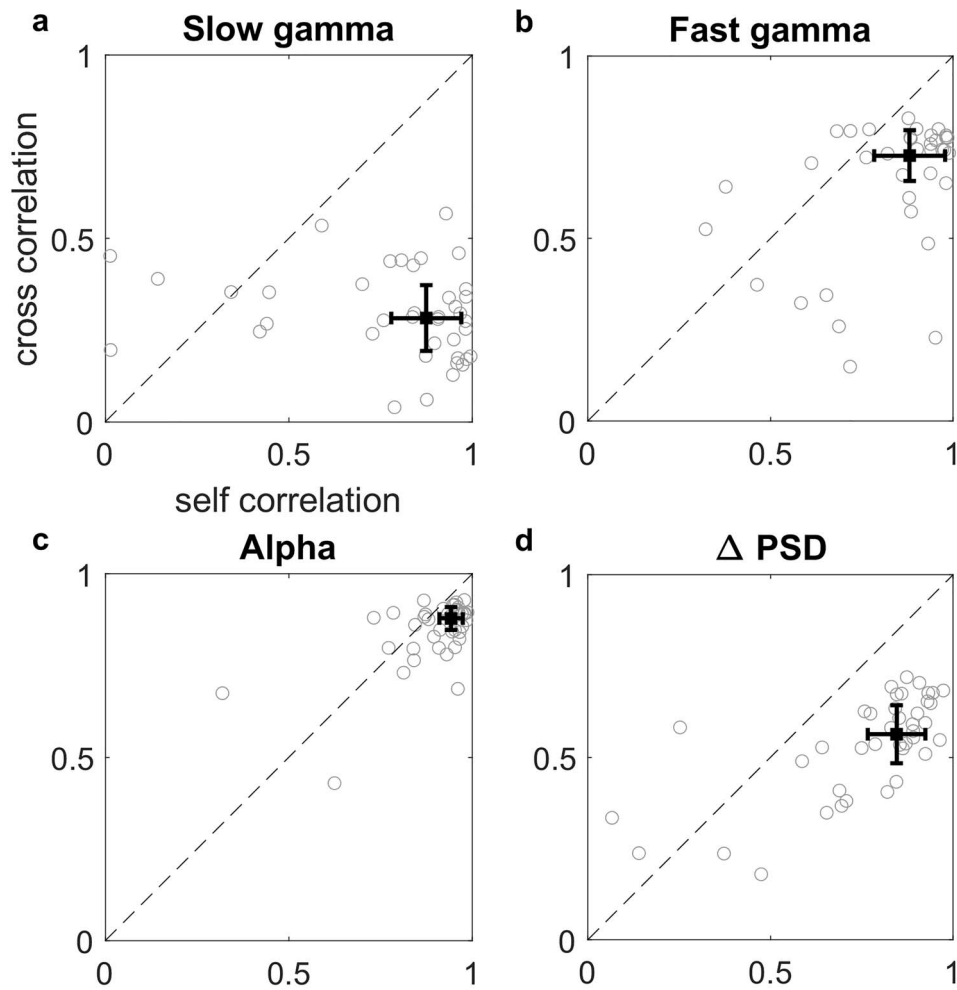
Figure 6. Change in PSD profile and temporal profiles of band powers are more correlated within self-pairs than cross-pairs. (A)–(C) Correlation between the band power time series traces (shown in Fig. 3) computed between 2 sessions of the same subject (self-pair) plotted against the median correlation between traces for all other subjects (cross-pair), for slow gamma (A), fast gamma, (B) and alpha (C) bands. (D) Same plot for the change in PSD traces shown in Figure 4. There are 40 data points in each plot. The median of these data is shown by the black box with error bars indicating median absolute deviation (MAD).

cross-pair correlation (0.56 ± 0.08; left-sided Wilcoxon rank sum test, $P = 3.36\text{e}{-}13$, $n = 40/1560$; Fig. 6d).

Feature-Based Subject Identification

From the data collected under "gamma" experiment, we computed subject distinctness measures, reflecting the separation between the self-pairs (40/1600) and cross-pairs (1560/1600), using spectral and temporal profiles of the signal. The separation was assessed within a classifier framework, based on 7 stimulus-induced features including band power differences (slow gamma, fast gamma, and alpha bands; 3 features), temporal profiles (3 features), and spectral profile (1 feature; Fig. 7; see Methods for details) using LDA. Feature weights were computed by 5-fold stratified training and was used for classifying test data.

Figure 7a shows the mean values of the 7 features for the cross and self-pairs. As expected, the band power differences were lower for self-pairs than cross-pairs (plots on the left). To use the correlation-based features in the same way to reflect dissimilarity, we subtracted

1 from the correlation values so that they were also lower for self-pairs than cross-pairs (right plots). To compare the distinctness of these features, and for data standardization, we z-scored each feature value across all the 1600 (40×40) pairs (Fig. 7b). All features had lower z-scores for self-pairs compared with cross-pairs, with the temporal profile of slow-gamma and the spectral profile having the largest separation in z-scores.

Next, we performed LDA on these z-scored data. Because data are z-scored, the weights of the classifier (shown in the legends of Fig. 7b) can be compared directly and reflect the importance of the features toward classification. These weights were also the highest for the temporal profile of slow-gamma and the spectral profile. Note that these weights are not always proportional to the difference in z-scores between the 2 classes, since some features may be highly correlated with others and may therefore carry redundant information (for example, temporal variation in fast gamma had a weight of almost zero).
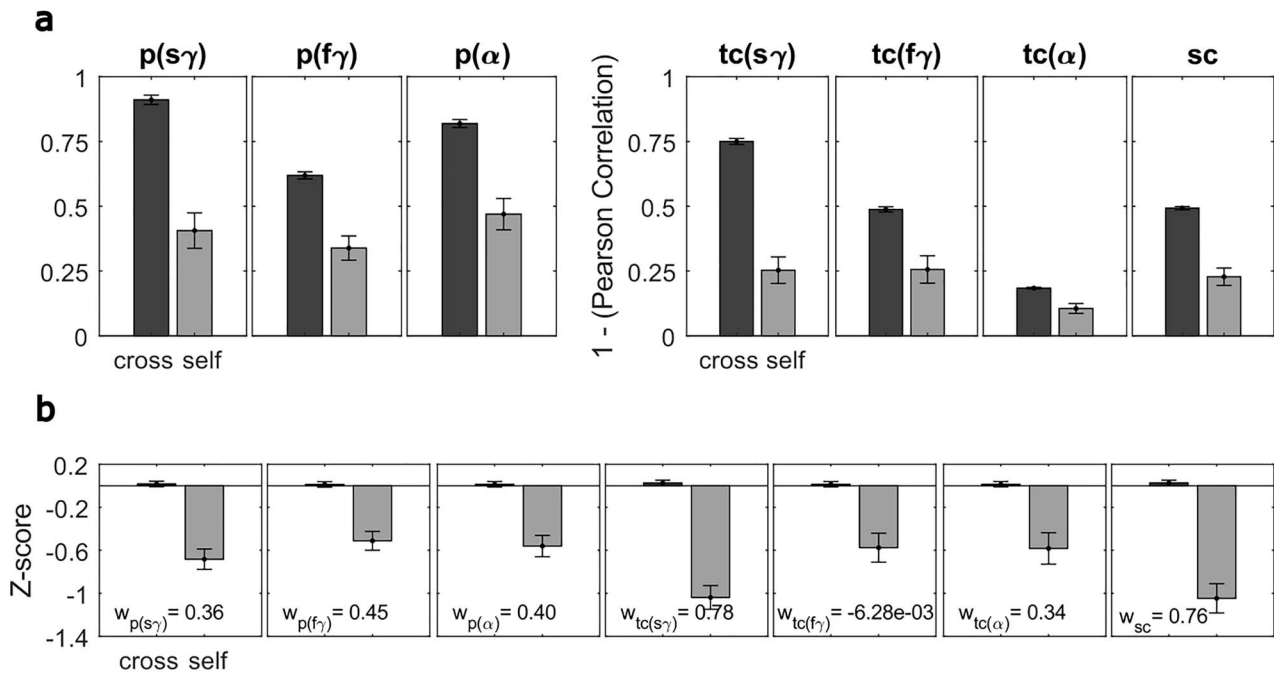
Figure 7. Temporal slow gamma and spectral correlations emerge as dominant features in classification. (A) The 7 features provided as inputs for the subject distinctness analysis are plotted separately for cross-pairs (cross individual comparisons) and self-pairs (same individual comparisons). Change in power (first 3 features) are in units of dB. Temporal and spectral correlations (other 4 features) are unitless. (B) Z-scored features are plotted in the same way as panel (A). The values mentioned in the figure insets denote the classifier weight of the respective feature in separating out self-pairs from the cross-pairs.

Further, we assessed the performance of each feature separately by calculating the AUC, which is a measure of the separation between the data in the 2 classes that does not depend on any explicit threshold (Fig. 8). Again, AUC values were highest for the temporal correlation of slow-gamma (0.84) and spectral correlation (0.83). Temporal correlation features remained dominant even when calculated over [0.25–0.75] s to exclude the transient activity. However, including all 7 features improved the AUC only marginally to 0.89, suggesting that these features had high redundancy. We also considered the baseline features (features obtained from resting state) including absolute average alpha power (1 feature; from the PSD in baseline period, Supplementary Fig. S3) and spectral profile (1 feature; correlation between the PSDs shown in Supplementary Fig. S3; see Methods for details) to determine the improvement in classification with inclusion of stimulus-induced features. Features obtained from resting state data also performed well (AUC = 0.82), with resting state alpha power (AUC = 0.73) performing better than spectral correlation of baseline PSDs (AUC = 0.67). The features in the baseline period were redundant with the stimulus features; adding these baseline features only improved the AUC to 0.90 (Fig. 8). These observations establish slow gamma temporal correlation and spectral correlation as the dominant features.

Discussion

We studied reliability of various spectral and temporal neural markers on visual stimulus-induced responses in EEG recorded from middle-aged/elderly (50–88 years) healthy human subjects, with a test–retest interval of about 1 year. Time frequency spectra were consistent across baseline and follow-up sessions for both males and females. Similar consistency was observed for temporal traces of alpha, gamma band powers as well as change in PSD profiles. We used ICC to measure reliability and found power in alpha, slow, and fast gamma to be significantly reliable. SSVEP power at 32 Hz were also found to be reliable. Temporal and spectral profiles were tested for reliability through Pearson's correlation between the profiles of same individual pairs ("self-pair correlation") and were found to be significantly higher than the correlations between cross individual pairs ("cross-pair correlation"). We trained a linear classifier to distinguish pairs of datasets as belonging to the same subject (self) versus different (cross) using 7 features including absolute difference of powers (3 features), temporal band power correlations (3), and spectral correlation (1). The feature weights as well as the separation in individual feature z-scores were highest for slow gamma temporal correlation and spectral correlation. We measured classifier performance using area under ROC curve and found that each of these 2 features also performed the best (AUC of ~0.84). Overall classification performance increased only modestly when all features were used (AUC of 0.89), suggesting high redundancy between features. Classification using only resting state (baseline) features was also high (AUC of 0.82 when resting alpha and baseline PSDs were considered) and combining resting state and stimulus-induced features yielded AUC of 0.90. Overall,
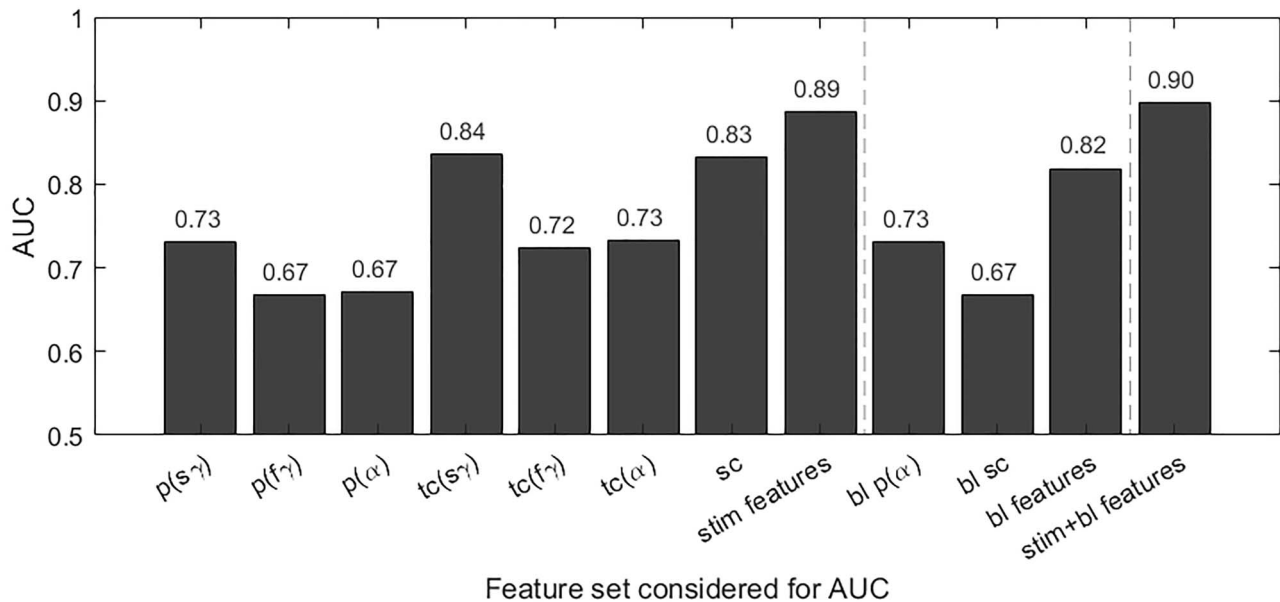
Figure 8. Features from stimulus period perform better than resting state features. The AUC values obtained using various classifiers, constructed based on the feature set listed on the *x*-axis. First, the performance of each of the stimulus-induced features is considered individually and subsequently collectively, and then the baseline (resting state) features are considered, first individually and subsequently collectively, and finally combined with the stimulus-induced features.

our results suggest that stimulus-induced spectral and temporal changes in various frequency bands are reliable and distinct, allowing their use in subject identification as well.

Comparison with Previous Studies

Reliability of Fast and Slow Gamma

Visually induced gamma oscillations in fast gamma range were shown to be reliable in human MEG with substantially higher ICC(C,1) of 0.89 (Hoogenboom et al. 2006; Tan et al. 2016) compared with our results (ICC(C,1) of 0.72 for fast gamma). However, these 2 studies had much lower test–retest period of 1–11 day (Tan et al. 2016) and 4 days–4 months (Hoogenboom et al. 2006), and the stimulus was a contracting circular sine wave grating. Hence, the low ICC in our case could be due to multiple reasons, such as the stimulus, recording modality and the test–retest period. Test–retest interval seems to be an important factor, since McCusker et al. (2021) who studied induced gamma in MEG after a period of up to 3 years reported much lower agreement ICC(A,1) of 0.55 (McCusker et al. 2021). Apart from the duration, the lower values could be due to the ICC type used, as the agreement type ICCs are usually lower than the consistency and classical definitions in presence of systematic bias in the data (although we did not find this in our data). Moreover, it could be due to the use of a checkerboard stimulus. Another human MEG study showed test–retest reliability in induced gamma oscillations with 1 week gap (Muthukumaraswamy et al. 2010), but they did not report ICC values.

This is the first time that the reliability of induced slow gamma oscillations was addressed in human EEG with a gap of approximately a year. Slow gamma was observable in a single subject (S02) in the study by Tan et al. (2016) in human MEG, but they do not follow-up on it further. They used many stimuli with differences in spatial structure (circular or Cartesian grating), temporal structure (moving or stationary), and size (large or small), however, only some of which produced slow gamma. We have previously shown that slow gamma oscillations are induced in human EEG with stationary Cartesian gratings of large size (Murty et al. 2018) and hence we employed similar stimulus parameters for this study.

Reliability of Alpha Band

Corsi-Cabrera et al. (2007) found less intersession reliability in absolute alpha band power with eyes open compared with beta band. They attributed less intersession reliability to the higher reactivity of alpha to external stimulation modifying the subjects' internal state. Several studies have compared absolute versus relative power (percentage value of power within the band w.r.t total power across all the bands), with mixed results. Some studies found relative alpha power to be more reliable than absolute power (John et al. 1983; Kondacs and Szabó 1999), others found comparable reliability (Gasser et al. 1985; Salinsky et al. 1991), whereas a recent study (McCusker et al. 2021) reported more robust reliability of absolute alpha activity in resting state than in the stimulus period (alpha desynchronization). In our study, relative power (ICC(1) = 0.81) measures were more reliable than absolute power (ICC(1) = 0.76) in the alpha band, within the stimulus period. Moreover, in line with the study by McCusker et al. (2021), the resting state alpha reliability of both the relative power (ICC(1) = 0.83) and absolute power (ICC(1) = 0.78) were higher than the stimulus period. Apart from alpha power, alpha peak

frequency also has been investigated and found to be test–retest reliable (Maltez et al. 2004; Näpflin et al. 2007, 2008; Grandy et al. 2013).

Reliability of Other Signals

Test–retest reliability has been investigated in signals such as event-related synchronization and desynchronization (ERS/ERD) (Neuper et al. 2005), theta and alpha oscillations during a working memory task (McEvoy et al. 2000), functional EEG network characteristics (Velde et al., 2019) and evoked oscillations with visual stimuli (Keil et al. 2003; Fründ et al. 2007). Although the degree of reliability is difficult to compare across studies, all measures have shown some degree of test–retest reliability. Similarly, spectral features like the spectral exponent (slope of the PSD in log–log coordinates) have been shown to convey test–retest reliability (Demuru and Fraschini 2020).

Although, to our knowledge, reliability of SSVEPs in gamma range has not been studied, there are several studies on its auditory equivalent, the auditory steady-state response (ASSR) (McFadden et al. 2014; Legget et al. 2017; Hirano et al. 2020). Interestingly, ASSRs have been shown to be abnormal in patients with autism (Wilson et al. 2007) and schizophrenia (Brenner et al. 2003). Hence, the aforementioned reliability studies have validated the use of ASSRs as a neuropsychiatric biomarker, including those employing MEG recordings (Tan et al. 2015). A recent study further showed ASSRs to be reliable in schizophrenia patients (Roach et al. 2019).

Subject Distinctness

Näpflin et al. (2007) used alpha power, center frequency, and spectral correlation (in [2–32 Hz]) for data collected under eyes closed and eyes open paradigm to separate out same individual pair from interindividual pair using a Generalized Linear Model (GLM) and reported 88% sensitivity or probability of detection with 40 subjects. In their study, GLM was fitted for each person separately. They also reported remarkable performance by the spectral shape feature. Another study used short time variability of the EEG spectral pattern for subject recognition using automatic speaker identification method (Stassen 1980). This study used supervised clustering with a cluster for each subject. We instead formulated a classifier over all the subjects and measured performance using AUC. Kondacs and Szabo (1999) found lower intraindividual than interindividual variability mainly in alpha power (7.5–12.5 Hz) and alpha mean frequency in eyes closed condition. We chose spectral shape feature (spectral correlation) (Fig. 7a) and alpha power during resting state (Supplementary Fig. S4). Computation of alpha mean frequency was not informative in our case since the spectral resolution was 2 Hz (due to analysis duration of 0.5 s), giving rise to only 3 bins in alpha range (8–12 Hz). Hence, we did not consider alpha mean frequency feature in this study.

Dominance of Spectral Profile and Slow Gamma Temporal Correlation and in Classification

High performance of spectral profile feature is not unexpected, since this feature captures the changes in power at all frequencies including alpha and gamma bands, as well as features such as the shape of the PSD, and has been previously shown to perform well in classification tasks (Näpflin et al. 2007). More surprising was the high performance of the temporal profile of slow gamma. As shown in Figure 3, first column, slow gamma temporal profile was more variable across subjects than other measures such as fast gamma. We have previously shown that unlike fast gamma which is induced soon after stimulus onset and then is subsequently maintained throughout the stimulus duration, slow gamma builds up slowly over time (Murty et al. 2018). On the other hand, the transient broadband component is high early after stimulus onset (Figs 1 and 2). Depending on the ratio of transient and steady-state components in the temporal profile, slow gamma temporal profile showed higher variability across subjects than fast gamma, which generally was highest after stimulus onset and then decayed (Fig. 3).

Another reason could be the nature of slow gamma itself. Slow gamma has larger functional spread, that is, larger coherent neural population in comparison with fast gamma, as shown in monkey LFP (Murty et al. 2018). Such large-scale synchrony might contribute to its relatively higher reliability. The dominance of slow gamma over fast gamma also has important practical considerations. In EEG, higher frequencies (>30 Hz) often get corrupted by a high noise floor (as observed in a few cases in our data as well; see Supplementary Fig. S3). In addition, the line noise at 50 or 60 Hz is within the fast gamma range. These factors impede accurate estimation of fast gamma in EEG (which could have contributed to a lower performance in our data as well). Slow gamma does not suffer from these practical limitations. Therefore, using static, full field gratings that induce strong slow gamma could be a convenient way to obtain a richer spectral signature that is also more reliable with age than fast gamma (Murty et al. 2020) and is equally compromised in AD as fast gamma (Murty et al. 2021).

Factors that Influence Reliability

The reliability of gamma oscillations could depend on the stability of inhibitory interneuron networks that are thought to generate these rhythms (Buzsáki and Wang 2012). In addition, oscillations recorded on the scalp could be influenced by various subject dependent parameters like head geometry, scalp conductivity, and other low-level factors. Further, some studies have shown cortical thickness to be correlated with gamma peak frequency, but not with power values (Muthukumaraswamy et al. 2010; Gaetz et al. 2012). Auditory steady-state activity (ASSR) at 40 Hz was also shown to be positively associated with left superior temporal gyrus cortical thickness in healthy controls (Edgar et al. 2014). Since

these external factors may also change with time, the reliability of the recorded signal reflects both the stability of the sources and these confounding factors. Some of these factors such as cortical folding (Schultz et al. 2013), cortical thickness (Gaetz et al. 2012), and surface area (Herting et al. 2015) can be estimated using subject-specific magnetic resonance imaging (MRI). Moreover, reliability estimates could be improved by increasing the signal to noise ratio by using accurate head models allowing us to go from the sensor to source space (Hoogenboom et al. 2006; Muthukumaraswamy et al. 2010; Martín-Buro et al. 2016; Tan et al. 2016).

Supplementary Material

Supplementary material can be found at *Cerebral Cortex Communications* online.

Notes

Conflict of Interest: None declared.

Funding

References

Bokil H, Andrews P, Kulkarni JE, Mehta S, Mitra PP. 2010. Chronux: a platform for analyzing neural signals. *J Neurosci Methods*. **192**: 146–151.

Brenner CA, Sporns O, Lysaker PH, O'Donnell BF. 2003. EEG synchronization to modulated auditory tones in schizophrenia, schizoaffective disorder, and schizotypal personality disorder. *Am J Psychiatry*. **160**:2238–2240.

Buzsáki G, Wang X-J. 2012. Mechanisms of gamma oscillations. *Annu Rev Neurosci*. **35**:203–225.

Corsi-Cabrera M, Galindo-Vilchis L, del-Río-Portilla Y, Arce C, Ramos-Loyo J. 2007. Within-subject reliability and inter-session stability of EEG power and coherent activity in women evaluated monthly over nine months. *Clin Neurophysiol*. **118**:9–21.

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

Demuru M, Fraschini M. 2020. EEG fingerprinting: subject-specific signature based on the aperiodic component of power spectrum. *Comput Biol Med*. **120**:103748.

Edgar JC, Chen Y-H, Lanza M, Howell B, Chow VY, Heiken K, Liu S, Wootton C, Hunter MA, Huang M, et al. 2014. Cortical thickness as a contributor to abnormal oscillations in schizophrenia? *NeuroImage Clin*. **4**:122–129.

Fries P, Reynolds JH, Rorie AE, Desimone R. 2001. Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*. **291**:1560–1563.

Fründ I, Schadow J, Busch NA, Körner U, Herrmann CS. 2007. Evoked γ oscillations in human scalp EEG are test–retest reliable. *Clin Neurophysiol*. **118**:221–227.

Gaetz W, Roberts TPL, Singh KD, Muthukumaraswamy SD. 2012. Functional and structural correlates of the aging brain: relating visual cortex (V1) gamma band responses to age-related structural change. *Hum Brain Mapp*. **33**:2035–2046.

Garcia-Rill E, D'Onofrio S, Mahaffey SC, Bisagno V, Urbano FJ. 2019. Bottom-up gamma and bipolar disorder, clinical and neuroepigenetic implications. *Bipolar Disord*. **21**:108–116.

Gasser T, Bächer P, Steinberg H. 1985. Test-retest reliability of spectral parameters of the EEG. *Electroencephalogr Clin Neurophysiol*. **60**: 312–319.

Gieselmann MA, Thiele A. 2008. Comparison of spatial integration and surround suppression characteristics in spiking activity and the local field potential in macaque V1. *Eur J Neurosci*. **28**: 447–459.

Grandy TH, Werkle-Bergner M, Chicherio C, Schmiedek F, Lövdén M, Lindenberger U. 2013. Peak individual alpha frequency qualifies as a stable neurophysiological trait marker in healthy younger and older adults. *Psychophysiology*. **50**:570–582.

van der Velde B, Haartsen R, Kemner C. 2019. Test-retest reliability of EEG network characteristics in infants. *Brain Behav*. **9**:e01269.

Hassler U, Trujillo Barreto N, Gruber T. 2011. Induced gamma band responses in human EEG after the control of miniature saccadic artifacts. *Neuroimage*. **57**:1411–1421.

Henrie JA, Shapley R. 2005. LFP power spectra in V1 cortex: the graded effect of stimulus contrast. *J Neurophysiol*. **94**: 479–490.

Herrmann CS, Munk MHJ, Engel AK. 2004. Cognitive functions of gamma-band activity: memory match and utilization. *Trends Cogn Sci*. **8**:347–355.

Herting MM, Gautam P, Spielberg JM, Dahl RE, Sowell ER. 2015. A longitudinal study: changes in cortical thickness and surface area during pubertal maturation. *PLoS One*. **10**:e0119774.

Hirano Y, Nakamura I, Tamura S, Onitsuka T. 2020. Long-term test-retest reliability of auditory gamma oscillations between different clinical EEG systems. *Front Psych*. **11**Available at::876. doi: https://doi.org/10.3389/fpsyt.2020.00876 [last accessed 8 June 2021].

Hoogenboom N, Schoffelen J-M, Oostenveld R, Parkes LM, Fries P. 2006. Localizing human visual gamma-band activity in frequency, time and space. *Neuroimage*. **29**:764–773.

Iaccarino HF, Singer AC, Martorell AJ, Rudenko A, Gao F, Gillingham TZ, Tsai LH. 2016. Gamma frequency entrainment attenuates amyloid load and modifies microglia. *Nature*. **540**:230–235. https://doi.org/10.1038/nature20587.

Jensen O, Kaiser J, Lachaux J-P. 2007. Human gamma-frequency oscillations associated with attention and memory. *Trends Neurosci*. **30**:317–324.

Jia X, Xing D, Kohn A. 2013. No consistent relationship between gamma power and peak frequency in macaque primary visual cortex. *J Neurosci*. **33**:17–25.

John ER, Prichep L, Ahn H, Easton P, Fridman J, Kaye H. 1983. Neurometric evaluation of cognitive dysfunctions and neurological disorders in children. *Prog Neurobiol*. **21**:239–290.

Keil A, Stolarova M, Heim S, Gruber T, Müller MM. 2003. Temporal stability of high-frequency brain oscillations in the human EEG. *Brain Topogr*. **16**:101–110.

Kitchigina VF. 2018. Alterations of coherent theta and gamma network oscillations as an early biomarker of temporal lobe epilepsy and Alzheimer's disease. *Front Integr Neurosci*. **12**:36.

Kondacs A, Szabó M. 1999. Long-term intra-individual variability of the background EEG in normals. *Clin Neurophysiol*. **110**:1708–1716.

Legget KT, Hild AK, Steinmetz SE, Simon ST, Rojas DC. 2017. MEG and EEG demonstrate similar test-retest reliability of the 40Hz auditory steady-state response. *Int J Psychophysiol*. **114**: 16–23.

Maltez J, Hyllienmark L, Nikulin VV, Brismar T. 2004. Time course and variability of power in different frequency bands of EEG during resting conditions. *Neurophysiol Clin Neurophysiol*. **34**: 195–202.

Martín-Buro MC, Garcés P, Maestú F. 2016. Test-retest reliability of resting-state magnetoencephalography power in sensor and source space. *Hum Brain Mapp*. **37**:179–190.

McCusker MC, Lew BJ, Wilson TW. 2021. Three-year reliability of MEG visual and somatosensory responses. *Cereb Cortex*. **31**:2534–2548.

McEvoy LK, Smith ME, Gevins A. 2000. Test–retest reliability of cognitive EEG. *Clin Neurophysiol*. **111**:457–463.

McFadden KL, Steinmetz SE, Carroll AM, Simon ST, Wallace A, Rojas DC. 2014. Test-retest reliability of the 40 Hz EEG auditory steady-state response. *PLoS One*. **9**:e85748.

McGraw KO, Wong SP. 1996. Forming inferences about some intraclass correlation coefficients. *Psychol Methods*. **1**:30.

Murty DVPS, Shirhatti V, Ravishankar P, Ray S. 2018. Large visual stimuli induce two distinct gamma oscillations in primate visual cortex. *J Neurosci*. **38**:2730–2744.

Murty DVPS, Manikandan K, Kumar WS, Ramesh RG, Purokayastha S, Javali M, Rao NP, Ray S. 2020. Gamma oscillations weaken with age in healthy elderly in human EEG. *Neuroimage*. **215**: 116826.

Murty DV, Manikandan K, Kumar WS, Ramesh RG, Purokayastha S, Nagendra B, ML A, Balakrishnan A, Javali M, Rao NP, et al. 2021. Stimulus-induced gamma rhythms are weaker in human elderly with mild cognitive impairment and Alzheimer's disease. *Elife*. **10**:e61666.

Muthukumaraswamy SD, Singh KD, Swettenham JB, Jones DK. 2010. Visual gamma oscillations and evoked responses: variability, repeatability and structural MRI correlates. *Neuroimage*. **49**: 3349–3357.

Näpflin M, Wildi M, Sarnthein J. 2007. Test–retest reliability of resting EEG spectra validates a statistical signature of persons. *Clin Neurophysiol*. **118**:2519–2524.

Näpflin M, Wildi M, Sarnthein J. 2008. Test–retest reliability of EEG spectra during a working memory task. *Neuroimage*. **43**:687–693.

Neuper C, Grabner RH, Fink A, Neubauer AC. 2005. Long-term stability and consistency of EEG event-related (de-)synchronization across different cognitive tasks. *Clin Neurophysiol*. **116**:1681–1694.

Ray S, Maunsell JHR. 2010. Differences in gamma frequencies across visual cortex restrict their possible use in computation. *Neuron*. **67**:885–896.

Roach BJ, D'Souza DC, Ford JM, Mathalon DH. 2019. Test-retest reliability of time-frequency measures of auditory steady-state responses in patients with schizophrenia and healthy controls. *NeuroImage Clin*. **23**:101878.

Salarian A. 2016. *Intraclass correlation coefficient (ICC)*. MATLAB cent file Exch Available at: https://academic.oup.com/cercor/article/31/5/2534/6042336.

Salinsky MC, Oken BS, Morehead L. 1991. Test-retest reliability in EEG frequency analysis. *Electroencephalogr Clin Neurophysiol*. **79**: 382–392.

Sarnthein J, Andersson M, Zimmermann MB, Zumsteg D. 2009. High test–retest reliability of checkerboard reversal visual evoked potentials (VEP) over 8 months. *Clin Neurophysiol*. **120**:1835–1840.

Schultz CC, Wagner G, Koch K, Gaser C, Roebel M, Schachtzabel C, Nenadic I, Reichenbach JR, Sauer H, Schlösser RGM. 2013. The visual cortex in schizophrenia: alterations of gyrification rather than cortical thickness—a combined cortical shape analysis. *Brain Struct Funct*. **218**:51–58.

Shrout PE, Fleiss JL. 1979. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. **86**:420–428.

Stam CJ, van Cappellen van Walsum AM, Pijnenburg YAL, Berendse HW, de Munck JC, Scheltens P, van Dijk BW. 2002. Generalized synchronization of MEG recordings in Alzheimer's disease: evidence for involvement of the gamma band. *J Clin Neurophysiol*. **19**: 562–574.

Stassen HH. 1980. Computerized recognition of persons by EEG spectral patterns. *Electroencephalogr Clin Neurophysiol*. **49**:190–194.

Tan H-RM, Gross J, Uhlhaas PJ. 2015. MEG—measured auditory steady-state oscillations show high test–retest reliability: a sensor and source-space analysis. *Neuroimage*. **122**:417–426.

Tan H-RM, Gross J, Uhlhaas PJ. 2016. MEG sensor and source measures of visually induced gamma-band oscillations are highly reliable. *Neuroimage*. **137**:34–44.

Thomas KP, Vinod AP. 2018. EEG-based biometric authentication using gamma band power during rest state. *Circuits Syst Signal Process*. **37**:277–289.

Uhlhaas PJ, Singer W. 2010. Abnormal neural oscillations and synchrony in schizophrenia. *Nat Rev Neurosci*. **11**:100–113.

Uhlhaas PJ, Singer W. 2012. Neuronal dynamics and neuropsychiatric disorders: toward a translational paradigm for dysfunctional large-scale networks. *Neuron*. **75**:963–980.

Vázquez-Marrufo M, Galvao-Carmona A, Benítez Lugo ML, Ruíz-Peña JL, Borges Guerra M, Izquierdo Ayuso G. 2017. Retest reliability of individual alpha ERD topography assessed by human electroencephalography. *PLoS One*. **12**:e0187244.

Vialatte F-B, Maurice M, Dauwels J, Cichocki A. 2010. Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. *Prog Neurobiol*. **90**:418–438.

Wilson TW, Rojas DC, Reite ML, Teale PD, Rogers SJ. 2007. Children and adolescents with autism exhibit reduced MEG steady-state gamma responses. *Biol Psychiatry*. **62**:192–197.