HHS Public Access

Author manuscript

Nat Med. Author manuscript; available in PMC 2018 May 06.

Published in final edited form as:

Nat Med. 2017 December ; 23(12): 1474–1480. doi:10.1038/nm.4433.

Selective neuronal lapses precede human cognitive lapses following sleep deprivation

Yuval Nir[1], **Thomas Andrillon**[2,3,4], **Amit Marmelshtein**[1], **Nanthia Suthana**[4], **Chiara Cirelli**[4], **Giulio Tononi**[4,*], and **Itzhak Fried**[5,6,*]

[1]Department of Physiology and Pharmacology, Sackler School of Medicine, and Sagol School of Neuroscience, Tel Aviv University, Tel Aviv 69978, Israel

[2]Laboratoire de Sciences Cognitives et Psycholinguistique (UMR8554), EHESS/CNRS/ENS-DEC, 75005 Paris, France

[3]Ecole Doctorale Cerveau Cognition Comportement, ENS/EHESS/ParisVI/ParisV, 75005 Paris, France

[4]Department of Psychiatry, University of Wisconsin-Madison, 6001 Research Park Blvd, Madison, WI, 53719, USA

[5]Departments of Neurosurgery, Psychiatry, and Biobehavioral Sciences, David Geffen School of Medicine and Semel Institute For Neuroscience and Human Behavior, UCLA, 710 Westwood Plaza, Los Angeles, CA 90095, USA

[6]Functional Neurosurgery Unit, Tel Aviv Medical Center and Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

Abstract

Sleep deprivation (SD) is a major source of morbidity with widespread health effects including increased risks of hypertension, diabetes, obesity, heart attack, and stroke[1]. Moreover, SD brings about vehicle accidents and medical errors[2–4], and is therefore an urgent topic of investigation. During SD, homeostatic and circadian processes interact to build up sleep pressure[5] that results in slow behavioral performance (cognitive lapses) typically attributed to attentional thalamic and fronto-parietal circuits[6–14], but the underlying mechanisms remain unclear[3,15]. Recently, it was found in human electroencephalogram (EEG)[16,17] and in the local field potential (LFP) of non-human primates[18] and rodents[19] that during SD, regional 'sleep-like' slow/theta waves co-occur with impaired behavioral performance during wakefulness. Here we used intracranial electrodes to record single-neuron and LFP activities in human neurosurgical patients performing a face/non-

*Equal senior authorship

Author contributions. Y.N. C.C. G.T. and I.F. conceived and designed research, I.F. performed surgeries, Y.N. and N.S. collected data, Y.N. T.A. and A.M. analyzed data, Y.N. T.A. C.C. G.T. and I.F. wrote the manuscript. All authors provided ongoing critical review of results and commented on the manuscript.

Competing Financial Interests Statement. None to declare.

Data availability. The data that support the findings of this study are available from the corresponding author upon reasonable request.

face categorization psychomotor vigilance task (PVT)[20–24] in multiple experimental sessions, including after full-night SD. We find that just before cognitive lapses, selective spiking responses of individual neurons in the medial temporal lobe (MTL) are attenuated, delayed, and lengthened. These 'neuronal lapses' are evident on a trial-by-trial basis comparing the slowest behavioral PVT reaction times to the fastest. Furthermore, during cognitive lapses LFPs exhibit a relative local increase in slow/theta activity that is correlated with degraded single-neuron responses and with baseline theta activity. Our results show that cognitive lapses involve local state-dependent changes in neuronal activity already in the MTL.

To study the neural underpinning of SD in the human brain, we investigated selective responses of individual neurons and how such responses change upon behavioral lapses. Twelve patients with pharmacologically intractable epilepsy, who were undergoing depth electrode monitoring to identify seizure foci for potential neurosurgical treatment, performed a face/non-face categorization variant of the PVT in 31 experimental sessions (Figure 1a; see Supp. Table 1 for additional details). Each session included 2*12-minute blocks, where 6 images of famous people, familiar landmarks and animals were presented (24 trials each) for 200ms with long unpredictable inter-stimulus-intervals (2–8sec) as participants performed a face/non-face categorization task. In four individuals, pairs of PVT sessions were conducted before and after full-night SD conducted for clinical purposes (time spent awake after SD = 24.1 ±1.6 hours), thereby providing a unique opportunity to examine the effects of SD on behavior and underlying activity of individual neurons. In two individuals, we were also able to acquire data during four PVT sessions conducted before and after normal sleep to address possible circadian and learning effects. Subjects performed the task successfully and accurately (percent correct = 94.1 ± 1.9%, no response = 2.9 ± 1.1%, mean ± SEM across 31 sessions). In subsequent analyses, we focus only on correct responses to maximize the chances that changes in behavior or neuronal activity in some trials were driven by the subjects' internal state, rather than by an impoverished visual stimulus.

The distribution of behavioral reaction times (RTs) during PVT experiments was best fit by an ExGaussian function, representing a mixture of a normal distribution (standard RTs) and an exponential distribution for slow RTs ('right-tail'), with large variability between sessions and between participants in the predominance of the exponential component (Figure 1b & Supp. Fig. 1), as has been previously observed in healthy individuals[21]. In each experimental session, 'fast trials' (fastest RTs) and 'slow trials' (slowest RTs) were defined according to the fitted ExGaussian distributions (Methods), with slow trials comprising 16.9 ± 1.0% of correct responses (mean ± SEM, n=31 sessions). We use the term 'cognitive lapses' throughout to refer to these slow trials characterized by delayed behavioral responses (rather than a complete absence of response), as customary in the SD PVT literature[20–24]. Other strategies for defining cognitive lapses, such as selecting the slowest/fastest 5–10% of trials yielded similar results (not shown). Figure 1b illustrates how sleep deprivation (SD) altered the distribution of RTs. The reciprocal of behavioral RTs (1/RT) is a sensitive marker of slower performance after sleep loss[20] and was decreased after SD (−21.3%, Figure 1c). The parameter tau (τ), denoting the exponential decay component of the ExGaussian function, increased from 122.9 ± 41.7 ms (mean ± SD) to 306 ± 237 ms (+128% ± 79, Figure 1c) whereas the parameter mu (μ), denoting the mean of the normal distribution, was only

modestly increased (+14.4% ± 8.1, mean ± SEM), indicating that SD exerted its greatest effect on cognitive lapses, as previously reported[23]. We also observed a 'time on task' effect[23] whereby the frequency of cognitive lapses increased with time spent performing the task (Supp. Fig. 2).

Given that SD had a marked effect on cognitive lapses, we next examined whether time spent awake (TSA) before each session could predict the extent of slow trials across the entire dataset (including sessions not conducted after SD). ANOVA analysis using TSA and the number of times a participant had performed the task (training effect) as between-session factors revealed that only TSA significantly predicted slow trials ($F=4.3$, $p=0.047$, mean RTs for cognitive lapses across sessions, training effect and interaction: $p>0.05$) and the average RTs ($F=4.4$, $p=0.046$; other $p>0.05$). None of the variables analyzed could significantly predict mean RTs for 'fast trials' across sessions. Time of day (circadian effect) did not correlate with RTs (fast or slow: all $p>0.5$); however, the lack of significant circadian effects here could stem from variability in the precise hours when sessions were conducted (Supp. Table 1). By contrast, accuracy was not affected by TSA ($F=0.03$, $p=0.86$). A relation between TSA and slow trials was likewise evident when conducting ANOVA on mu (μ), sigma (σ) and tau (τ) (the parameters of the fitted ExGaussian distribution corresponding to mean, left-tail and right-tail, respectively): a significant effect was found between TAS and tau (τ) ($F=4.7$, $p=0.04$; see Fig. 1d), but no effect of TSA on mu (μ)($F=0.54$, $p=0.47$) or sigma (σ)($F=0.69$, $p=0.41$) verifying the specific relationship between TSA and cognitive lapses during slow trials. We did not find an effect of repeated sessions (training effect) on performance, in accordance with the literature[23]. In addition, subjective sleepiness (Methods) was significantly correlated with TSA (Spearman's $r=0.42$, $p<0.05$), but ANOVA did not reveal a significant relation between subjective sleepiness and slow trials ($F=1.5$ and 2.2 for slow trials' mean RT and tau (τ), respectively), replicating previous studies on the limitation of subjective sleepiness estimates[25]. Altogether, behavioral data show that TSA was the dominant factor influencing performance on the task, primarily increasing the occurrence of cognitive lapses.

Next, we examined the neuronal activity evoked by the stimuli used in the face/non-face categorization PVT task. Fig. 2a–b provide a schematic of the intracranial electrodes and the 104 brain regions monitored in the study. We focused on comparing neuronal activity during cognitive lapses with activity during 'fast trials' (lowest RTs) in the same experimental sessions. Importantly, such 'within-session' comparisons minimize confounds of increased epileptogenic activity after SD. The visual stimuli used in the face/non-face categorization PVT paradigm elicited robust responses in individual neurons (Fig. 2c, 1481 units recorded in total), especially in the MTL but occasionally also in cingulate cortex, with variability in the precise intensity, selectivity, and latency of responses across individual neurons. Whenever possible, images were chosen based on prior screening sessions to maximize the likelihood of eliciting responses in the recorded neurons. Of 611 neurons recorded in the MTL, 106 (17%) responded significantly to at least one stimulus ('responsive MTL neurons', Methods). When pooling the activity of all responsive neurons irrespective of brain region (n=162), an average response profile emerged consisting of increased firing rates 200–500ms after stimulus onset (Fig. 2d) with an orderly progression of temporal latencies from high-order visual cortex to hippocampus and frontal lobe (Fig. 2e). Robust

differences in response latencies along the visual-hippocampal hierarchy were also evident when quantifying the precise timing of responses detected in each trial separately (Methods, Supp. Figs. 3 & 4), in line with previous findings[26,27]. Single-unit spiking responses were highly selective and could not be observed when averaging the activity of neighboring neurons not categorized as responsive (Supp. Fig. 5). Crucially, the robust and highly selective profiles of single-neuron responses allowed studying the effects of cognitive lapses at the single-neuron level.

We examined the relationship between cognitive lapses and underlying neuronal activity by testing how the responses of the same neurons to the same physical stimulus may change as a function of behavioral performance ('fast trials' vs. 'slow trials' in the same session). Given the relation between cognitive lapses and TSA (Fig. 1D), we compared neuronal responses across all sessions (n=31), 15.3% of which were obtained before/after SD, and 71.4% in sessions when subjects were awake for > 12 hours. In individual neurons, cognitive lapses on slow trials were associated with weaker and delayed neuronal spiking discharges, and differences were particularly evident around 200–300ms following image onset (Fig. 3a). We proceeded to examine the average normalized response in fast vs. slow trials across the entire dataset. As can be seen (Fig 3b,c,d) slow trials were associated with attenuated, delayed and prolonged responses to identical stimuli (see also Supp. Fig. 6 for individual subject data and Supp. Fig. 7A for non-normalized PSTHs). A quantitative paired comparison between the response of each individual neuron in fast and slow trials (Fig 3e, Methods) revealed that response magnitude was attenuated by 17% ($z(376)= -3.05$, $p=0.0023$ via Wilcoxon signed-rank test). Response latency (detected in individual trials, Methods) was delayed by 27ms ± 6.9 (mean ± SEM; $z(376)=3.5$, $p=4.8*10^{-4}$ via Wilcoxon signed-rank test). Additional analysis quantifying response latency as firing above baseline in PSTHs yielded similar results (not shown). Response duration was increased by 52ms ± 19 (mean ± SEM; $z(376)=3.2$, $p=0.0012$ via Wilcoxon signed-rank test). Importantly, analysis of spiking activity in neighboring non-responsive neurons during the same trials did not reveal significantly different firing rates between fast and slow trials ($p=0.36$ via Wilcoxon signed-rank test). Thus, altered neuronal spiking activity during cognitive lapses was specific for responsive neurons and did not reflect a global reduction of activity at those times. Neuronal spiking responses were best locked to stimulus onset rather than to motor responses (Supp. Fig. 7). Correlation between the latency of MTL neuronal responses and RTs was also observed across all trials, without focusing a-priori on comparing fast versus slow trials (Supp. Fig. 8).

We also examined the responses to images in locally referenced LFPs recorded from the same MTL microwires where single-unit neuronal activity was observed (Fig. 4). An extensively-studied phenomenon is the robust increase in broadband LFP gamma power that occurs following sensory stimulation in multiple modalities. This LFP signal is linked to the neuronal spiking activity of local neuronal populations[28], typically co-occurring with a decrease in low-frequency power also termed 'desynchronization'[28–31]. The 'induced power' LFP response to pictures (Fig. 4a) was in line with these findings, consisting of an increase in broadband gamma power (>45Hz, 50–600 ms) and a decrease in slow/theta power (2–10 Hz, 300–700 ms, see Supp. Fig. 9 for examples of the LFP dynamics in single trials). LFP responses were selective, whereby some MTL microwires (n=270 channels in 31

sessions) showed a robust response (Fig. 4a, 'responsive channels') while other neighboring channels (n=198 channels in 31 sessions) did not show significant modulations (Fig. 4c, 'non-responsive'), despite high-quality signals that allowed isolation of neuronal units (Methods).

In responsive LFP channels, cognitive lapses during slow trials were associated with a weaker increase in gamma power (Fig. 4b,e; −19.1%, z(270)=2.72, p=0.006, Wilcoxon signed-rank test) and a weaker decrease in slow/theta power (Fig. 4b,f; −76.2%, z(270)= −5.2, $p=2*10^{-7}$, Wilcoxon signed-rank test). In contrast, no significant effects of cognitive lapses were observed in neighboring non-responsive MTL channels (Fig. 4d,e,f; gamma: z(198)= −0.57, p=0.57; theta: z(198)= −0.98, p=0.33; Wilcoxon signed-rank tests). In contrast to 'induced' power changes, power of the evoked (average) LFP at 2–10Hz was lower during cognitive lapses (Supp. Fig. 10) suggesting that induced power effects reflect changes in ongoing activity rather than changes in the stimulus-evoked event-related potential. Furthermore, during cognitive lapses the latency of spiking responses negatively correlated with LFP gamma power (Fig. 4g, r=−0.17, p=0.006) and positively correlated with LFP slow/theta power (Fig. 4h, r=0.22, $p=4.5*10^{-4}$). The significant coupling between the degree of degradation in LFP and neuronal spiking responses suggests that these effects are tightly linked manifestations of neuronal lapses in selective circuits engaged in the task. Whether cognitive lapses (and underlying neuronal activity) assessed after SD are *qualitatively* similar or different from sporadic slow responses occurring throughout wakefulness remains an open question for future studies[3,6].

Given the growing literature on increased theta (6–10Hz) power as a correlate of sleep pressure[16,19,32], we examined theta power during baseline intervals preceding stimulus onset (Methods, Supp. Figure 11). First, we established that theta power in MTL LFPs was indeed associated with sleep pressure and cognitive lapses. We found that baseline theta power was (a) significantly correlated with time spent awake (Supp. Figure 11a, r=0.26, $p<4.07*10^{-6}$), (b) was elevated after full-night SD (Supp. Figure 11b, $p<2.74*10^{-5}$, via Wilcoxon signed-rank test), and (c) higher before cognitive lapse trials (Supp. Figure 11c, p<0.0001 via Wilcoxon signed-rank test). Baseline theta power also exhibited a modest albeit highly significant correlation with the level of slow/theta (2–10Hz) power during the response interval (Supp. Figure 11d,e; r=0.05, $p<4*10^{-37}$), suggesting that baseline theta might influence the degraded LFP response during cognitive lapse trials (Fig. 4). Overall, ongoing theta activity is increased with sleep pressure, and its decreased attenuation during cognitive lapses may lead to an impoverished neuronal and cognitive responses.

Finally, we ruled out potential contribution of pathological epileptiform activity. First, we confirmed that all the main findings (degraded neuronal and LFP responses) hold when discarding all data collected in regions eventually declared as being within the seizure onset zone (SOZ; not shown). Second, we detected inter-ictal spikes (IISs) across the entire LFP dataset (N=1648) to test whether such events may occur more frequently around cognitive lapses (Methods & Supp. Fig. 12a,b). IISs were detected around few (5.0±0.23%) trials, and significantly more frequently within the SOZ than in other regions (2.7-fold increase, $p<10^{-48}$ via Mann-Whitney U-test, Supp. Fig. 12c), attesting to successful detection. However, cognitive lapses were not associated with increased IISs when considering all data

(Supp. Fig. 12d, p=0.46 via Wilcoxon signed-rank test, N=1533 channels) or when considering only MTL regions where the selective neuronal effects were observed (Supp. Fig. 12f, p=0.48 via Wilcoxon signed-rank test, N=619 channels). In fact, when considering only sessions after complete sleep deprivation we found a small but significant reduction in IISs around cognitive lapses (Supp. Fig. 12e, p=0.025 via Wilcoxon signed-rank test, N=186 channels), Thus, we could not reveal consistent or robust relation between IISs and cognitive lapses.

Altogether, these findings show that, in sleep deprived humans engaged in a visual categorization task, selective neuronal spiking responses to images are attenuated, delayed, and lengthened prior to cognitive lapses, and such MTL modulations in spiking activity are associated with selective weakened decrease in slow/theta power in responsive LFP channels. Thus, degraded neuronal activity is evident already at the perceptual stage, where responses of individual neurons in selected trials can predict subsequent cognitive lapses. The extent to which these effects are regionally specific remains unclear, but the current results establish that within MTL regions cognitive lapses specifically affect responsive circuits engaged in the task. Progressive delays in neuronal activity may further accumulate in downstream decision/motor regions during cognitive lapses, ultimately leading to slower behavior. In line with the biased competition model of selective attention[33], degraded sensory cortical activity during cognitive lapses may fail to elicit high-quality perceptual representations, and thus cannot be effectively fed forward to frontal lobe regions that ultimately determine behavior. It still remains unclear whether degraded MTL activity strictly reflects impaired bottom-up signaling or whether additional top-down attentional mechanisms are at play.

Brief periods of silence (so called OFF periods) accompanied by slow waves in field potentials are hallmarks of non-rapid-eye-movement (NREM) sleep in both animals[34] and humans[35], and are associated with behavioral immobility and unresponsiveness. Following SD, awake rats exhibit local “sleep-like” slow/theta waves and shorter OFF periods that are associated with degraded behavioral performance[19]. Given that we could only record few neurons simultaneously in each brain region and that OFF periods in wakefulness are short (~80ms), it was not possible to reliably determine if such brief OFF periods occur in the human brain undergoing sleep deprivation. However, we find that slow/theta activity, previously linked to sleepiness[16,17,19,32], was increased before and during cognitive lapses, and these changes were associated with degraded spike responses. Impaired spike responses are observed in individual neurons engaged in a cognitive task without concurrent changes in the firing of neighboring neurons, and these changes predict specific cognitive impairments in sleep-deprived humans. The tight relation between MTL activity and perception[36] suggests that visual recognition itself may slow down as a result of SD. PVT lapses are stochastic i.e., they are unpredictable moment to moment due to the influence of a random variable. The present findings suggest that degraded neuronal and LFP responses in the MTL do predict them to some extent, which pushes the formal cause of cognitive lapses back another step in the neurobiological chain of events. The mechanisms underlying local neuronal lapses remain to be determined, although it is likely that transient instability in the activity of neuromodulatory systems, including cholinergic and noradrenergic neurons, may play a role[37]. Indeed, instability in pupil size (tightly linked with central noradrenergic

activity[38]) is correlated with alertness[39] and the synaptic release of acetylcholine is transiently diminished during poor behavioral performance[40].

Online Methods

Subjects

Twelve patients (age 19–52 years, 5 females) with pharmacologically intractable epilepsy underwent monitoring with depth electrodes for seizure foci identification and potential surgical treatment[41]. Patients provided written informed consent prior to participation in the research study, under the approval of the Medical Institutional Review Board at the University of California, Los Angeles, USA. Electrode location was based only on clinical criteria, and Dr. Itzhak Fried performed all surgeries. For each subject, localization of the seizure onset zone was based on recordings during hospital monitoring, in combination with prior functional and anatomical neuroimaging.

Four patients participated in a full overnight sleep deprivation (SD) session to increase the propensity for interictal epileptiform discharges and seizures, to aid clinical diagnosis. During these nights, medical staff verified via video that patients remained awake, and entered their room whenever there were signs of sleepiness or eye closure.

Data acquisition

For each patient, 8 to 12 flexible polyurethane depth electrodes were placed in some of the following regions: hippocampus, amygdala, entorhinal cortex, parahippocampal gyrus, anterior fusiform gyrus, temporal gyrus, fusiform gyrus, temporo-occipital junction; anterior, middle and posterior cingulate; supplementary motor area, inferior frontal gyrus, orbitofrontal cortex, parietal cortex, posterior temporal cortex, and temporal-parietal cortex. Electrode location varied between patients based on their clinical profiles (see Supp. Table 1 for details). Electrode positions were verified using post-implant computed tomography (CT) co-registered with pre-implant magnetic resonance (MR) imaging using BrainLab stereotactic and localization software (www.brainlab.com)[42]. Each depth electrode terminated in a set of eight insulated 40-μm platinum-iridium microwires[41] (impedances 200 to 500 kΩ) protruding from the tip and located 4–5mm from the most distal macro-electrode contact (see Fig. 1d in [35]). Microwire signals were simultaneously recorded continuously (Cheetah Recording System; Neuralynx, Tucson, AZ for one patient; Neuroport Recording System; Blackrock, Salt Lake City, UT for the other 11 patients), sampled at 28 kHz (1 patient) or 30 kHz (11 patients), band-pass filtered in hardware between 1Hz and 9kHz, and referenced locally to a ninth non-insulated microwire ("LFP").

Unit identification and spike sorting

Units were identified using the 'wave_clus' software package [43] as described previously [35]: (i) extracellular microwire recordings were high-pass filtered above 300Hz, (ii) a 5 SD threshold above the median noise level was computed, (iii) detected events were clustered using superparamagnetic clustering, and categorized as noise, single- or multi-unit clusters. Classification of single- and multi-unit clusters was based on the consistency of action potential waveforms, and by the presence of a refractory period for single units, i.e. less than

1% of inter-spike-intervals (ISIs) within 3ms, as in [35]. Overall, 1481 units were identified (561 putative single units, 920 multi-unit clusters).

Face/non-face categorization PVT paradigm and subjective sleepiness ratings

Patients participated in 31 sessions of a visual face/non-face PVT paradigm (Fig. 1a). We used pictures of familiar people, familiar landmarks and animals (instead of detection of a "bull's-eye" target used in classical PVT) to elicit robust responses in MTL neurons. When possible, pictures were chosen based on their effectiveness in eliciting responses in the recorded neurons by means of a 'visual screening' experiment performed earlier that day [44]. Each session included either one (n=11) or two (n=20) 12-minute blocks. During each block, 4 face images and 2 non-face images (places or animals) were presented on a laptop computer for 200ms while subjects performed a face/non-face categorization task. Each picture was presented 24 times in a pseudo-randomized order (a total of 144 trials), with long pseudo-random inter-stimulus-intervals of 2–8s (uniform distribution) as in classical PVT designs [23]. Patients were instructed to press one of two buttons (for face vs. non-face) as quickly as possible. Subjective sleepiness was assessed at the beginning of each experimental session using a combination of the Stanford Sleepiness Scale (SSS) and a visual-analog rating of sleepiness that were combined to a single sleepiness score (Supp. Table 1).

Analysis of behavioral data

The distribution of reaction times (RTs) for trials with a correct answer to the face/non-face categorization task (95% of trials) was fit in each session separately by an ExGaussian function[45,46], representing a mixture of a normal distribution (standard RTs) and an exponential distribution modeling slow trials representing cognitive lapses (Supp. Fig. 1). The resulting parameters of fitting this model were mu (μ) & sigma (σ) (describing the Gaussian component) and tau (τ, describing the exponential component) and the effects of sleep pressure and subjective sleepiness could be quantitatively evaluated on behavioral performance in each session.

Cognitive lapses were defined as the trials contained in the exponential part of the distribution. To define the boundary between the Gaussian and Exponential parts, we simulated 10,000 values using the Gaussian component (mu (μ) and sigma (σ)) and computed the RT cutoff at 99.9% of the simulated values. RTs above this threshold value were therefore considered as belonging to the Exponential tail. In certain cases, the tail of the ExGaussian distribution contained a large proportion of all RTs. To avoid highly heterogeneous numbers of cognitive lapses across sessions, in those cases we limited cognitive lapses to a maximum of 20% of the slowest trials. "Fast" trials (trials with lowest RTs) were defined as the N fastest trials, N being the number of cognitive lapses for the same experimental session. Only correct trials were considered in this analysis.

Time on task effects (Supp. Fig. 2) were evaluated for statistical significance by dividing each session (n=31) into four equal parts according to the trial order. In each part, we computed the average response time for all correct trials as well as the proportion of fast trials and slow trials as defined above. To overcome inter-session variability and compare

RTs across sessions in this context, each RT (for each quartile and per session) was normalized by the value for the first quartile (100%). Supp. Fig. 2 shows the gradual increase in the mean RT and the proportion of slow trials in the course of face/non-face categorization PVT experimental blocks.

Analysis of epileptiform inter-ictal spikes (IIS)

To rule out potential contribution of epileptiform inter-ictal spikes (IIS) around cognitive lapses[47], we performed an automated detection of IIS in 10/12 patients in whom the seizure onset zone was determined clinically. IIS detection methodology followed our previous work[35] as follows: continuous LFP data for each microwire was band-pass filtered between 50 and 150 Hz (4^{th}-order Butterworth filter). The envelope of the band-pass filtered data was then extracted using the Hilbert transform. We extracted the mean and standard deviation of this high-frequency envelope for 10-s-long consecutive intervals. For each such interval, epochs longer than 5ms during which the envelope exceeded 8SD above the mean were marked as IIS. Examples of IIS detections are available in Supp. Fig. 12. We then marked, for each LFP channel separately, those trials that were associated with IISs within [−2, 4]sec window around trial onset.

Analysis of spiking activity

Single- and multi-unit clusters were evaluated for their response to each picture separately. Neurons were declared 'responsive' for a given image if their baseline firing rate exceeded 2 spikes/sec and if the mean firing rate at [200, 500]ms post-stimulus onset was significantly greater than its baseline ([−600, 0]ms) firing rate ($p<0.005$ via paired t-test). We also repeated the main analyses in the paper using different criteria for 'responsiveness' including $p<0.05$ and $p<0.001$, as well as not discarding neurons with baseline below 2 spikes/sec, and this did not affect the main results. Neurons were declared non-responsive for a given image when this p-value was above 0.2.

Automatic response detection and quantification: response latency and termination was determined by Poisson spike-train analysis as in [26,27]. In this procedure, the inter-spike intervals (ISIs) of a given unit are processed continuously over a [-600, 1000]ms window and the onset of a spike-train is detected based on its deviation from a Poisson process, (i.e. exponential distribution of ISIs), where the unit's baseline firing rate is used to estimate the distribution. A response was detected whenever a series of short ISIs represented a significant ($p<0.005$) deviation from this exponential distribution (examples in Supp. Fig. 3). Variations of the critical p-value defining significant deviations did not significantly alter subsequent results. Only responses with at least 3 spikes were further considered. For each trial we determined the presence of a response given these criteria and, whenever present, we defined its onset latency [response termination] as the time between image onset and the first [last] spike of the response (green and red dots in Supp. Fig. 3). Only latencies within the first 600ms were considered and the overall latency of that neuron was taken as the mean across all trials. Response duration was defined in each trial as the difference between response termination and response onset.

We then quantified (Fig. 3e) differences between fast trials and slow trials in terms of (a) response magnitude, defined as the mean firing rate over a [200, 500]ms window normalized by the [−600, 0]ms baseline activity for the same trials, (b) response latencies, and (c) response duration, the latter two were defined as explained in previous paragraph. These quantifications were performed separately in both responsive neurons (Fig. 3e) and non-responsive neurons (reported in text). To further compare these differences in responsive units with non-responsive units, we paired each of the conditions (pictures) for which there was a significant response (n=469 conditions in 162 neurons) with an equal number of data randomly selected from non-responsive units recorded simultaneously in the same brain region. Out of the 162 responsive units, 142 could be paired in such a manner and were included in Fig. 3e that shows, for responsive neurons, the comparison between fast and slow trials within each stimulus condition.

For visualization purposes only (Figure 3b,c,d), we opted to go beyond neuron-to-neuron variability in response timing and amplitude to highlight differences between fast and slow trials in the response to each stimulus. To this end, response PSTHs were aligned (x-axis) such that time zero was defined for each stimulus condition as follows. The response PSTH across all trials was computed and smoothed with Gaussian kernel (σ = 50ms), and the moments during the rise and fall slopes at which the response reached 0.3 of its peak magnitude were defined as 'beginning' (time zero, vertical lines in Figure 3b,c,d), and 'end' (diagonal lines in Figures 3b,c), respectively. Furthermore, response amplitude (color scales in Fig. 3b,c and y-axis in Fig 3d) was normalized by dividing the PSTHs of both fast or slow trials by one joint scalar being the maximum of the peak firing rate of the two PSTHs. Finally, Figure 3c shows a joint color representation of slow and fast trial PSTHs within each response (stimulus) where hue (green to orange) represents the difference between the normalized firing rate response in fast and slow trials. Brightness represents the magnitude of the normalized response for fast or slow, depending on which is the greatest.

Analysis of LFP responses

Apart for Supp. Figs. 8 & 12, all analysis of field potentials is restricted to MTL microwires. Field potentials obtained from locally-referenced microwire LFPs (Fig. 4) were segmented around each picture presentation, and a time-frequency decomposition in individual trials ('induced' power) was performed with the EEGlab toolbox[48]. An analysis of time-frequency dynamics in the LFP averaged across trials ('evoked' power) is also presented in Supp. Fig. 10. Artifact trials were rejected when amplitude in the band-passed signal ([0.1, 100]Hz) exceeded 1.5 times the median standard-deviation of the maximal values for all trials of the corresponding session and for the corresponding channel. Such trials (~8% of data) were excluded from further analyses. Sessions with less than five fast or slow trials were also excluded from further analysis.

Time-frequency decomposition was performed on the unfiltered data by applying a Fast-Fourier Transform and Hanning window tapering (padding ratio of 2). The power for each frequency and time was normalized by the pre-stimulus activity ([−1500, −500]ms) to compute the average (median) power modulation, expressed in decibels (dB). A second

baseline correction was performed across sessions ([−500, 0]ms, subtraction) to compute the power modulation for 'all trials', 'fast trials', or 'slow trials' (Fig. 4a,c).

LFP time-frequency "regions of interest" (ROIs, illustrated by rectangles in Fig. 4a,c) were defined according to the mean response for 'all trials'. ROIs included (i) a broadband increase in the gamma frequency range ([45–100]Hz, [50, 600]ms after stimulus onset) and (ii) a decrease in slow/theta power ([2, 10]Hz, [300–700]ms after stimulus onset). Power in these time-frequency ROIs were then quantified and statistically compared (Fig. 4e–h).

We restricted all analysis of LFP time-frequency dynamics to those 468 channels where spike sorting revealed single- or multi-unit clusters (to ensure good signal even in 'non-responsive' channels). LFP 'responsive channels' (Fig. 4a, N=270; about 3/5 of channels) were defined as those channels with significant increases in gamma power ([45–100]Hz, [50, 600]ms after stimulus onset, $p<0.1$ via t-test) OR significant decrease in slow/theta power ([2, 10]Hz, [300–700]ms after stimulus onset, $p<0.1$ via t-test) for all trials. LFP 'non-responsive channels (Fig 4b, N=198; about 2/5 of channels) were defined as those channels where spike sorting revealed neuronal clusters but with no significant increases in gamma power ([45–100]Hz, [50, 600]ms after stimulus onset, $p>0.1$ via t-test) AND no significant decrease in slow/theta power ([2, 10]Hz, [300–700]ms after stimulus onset, $p>0.1$ via t-test) for all trials. Noisy channels (<5%) with absolute gamma power modulations over 1dB in the baseline (<−500ms) were discarded.

For comparing LFP power modulations with the neuronal spiking responses (Fig. 4g,h) we extracted, for each responsive units within the MTL (and for each picture this unit responded to), its average response latency (Fig. 3d) for fast and slow trials. We then extracted the LFP from the same microwire, and computed the power in the gamma and slow/theta time-frequency ROIs during the same trials. Correlation between unit and LFP responses was assessed using the Spearman's method across pairs of units/LFP channels (n=255 pairs in 87 units across 21 sessions, Fig. 4g,h).

Analysis of baseline theta activity

Locally-referenced LFPs in MTL responsive channels (as above) during the pre-stimulus baseline periods ([−2000 0]ms) were segmented, and trials with artifacts (as above) or with inter-ictal spikes were excluded from further analysis. Baseline theta power (Supp. Figure 11a–d) was calculated as the percent of total power contained within the 6–10Hz band, by applying a Fast-Fourier Transform and Hamming window tapering and normalizing the power spectrum to a distribution (sum = 1). When comparing baseline ([−2000 0]ms) theta (6–10Hz) activity to post-stimulus ([300 700]ms) slow/theta (2–10Hz) power, as seen in Supp. Figure 11d/e, power values were calculated using the EEGlab toolbox[48] with parameters (e.g. FFT, tapering, normalization) as described above. Similar results (e.g. $R=0.061$, $p<10^{-30}$ for Supp. Figure 11e) were obtained when quantifying 6–10Hz power in the response interval (same frequencies as in baseline). The relation between baseline theta and slow/theta power in response intervals was further evaluated in individual LFP channels by computing the correlation coefficient ("R" in Supp. Figure 11f) between baseline and response across all trials, and examining the distribution of R values across all LFP channels.

Statistics

Error bars in all figures denote standard error of the mean (SEM = SD/$\sqrt{(n-1)}$, where n is the number of data points) unless otherwise stated. Student T-tests were performed after confirming normal distributions via Kolmogorov-Smirnov tests. Non-parametric Mann-Whitney U-tests and Wilcoxon signed-rank tests were used whenever normality was not confirmed. Correlations were examined using the non-parametric Spearman's rank correlation coefficient or Pearson's method for normally distributed pairs of variables. To examine the circadian influence (i.e. time of day) on RTs (see main text), we could not use an ANOVA (circular variable). Instead, we computed the correlation coefficient between the time of the day and RTs using the 'circ_corrcl' function in the Circular Statistics Toolbox for Matlab. A Life Sciences Reporting Summary for can be found here: ##

Code availability

Analysis was performed in Matlab using custom-developed analysis routines and using the publicly available software package EEGLAB.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the patients for their cooperation; M. Tran, E. Behnke, T. Fields, H. Gelbard-Sagiv and M. Sagiv for assistance with data acquisition; B. Salaz and N. Regev for administrative help. This work was supported by the Human Frontier Science Program (HFSP) Organization long-term fellowship (YN), the I-CORE Program of the Planning and Budgeting Committee and the Israel Science Foundation (grant no. 51/11, YN), the FP7 Marie Curie Career Integration Grant (YN), the Adelis Foundation (YN), by the Société Française de Recherche et Médecine du Sommeil (SFRMS, TA), ANR-10-LABX-0087 IEC (TA), ANR-10-IDEX-0001-02 PSL* (TA), by NIMH grant R01MH099231 (CC and GT), NINDS grant P01NS083514 (CC and GT), R01GM116916 (GT), and NIH R01 NS033221 (IF).

References

1. Colten, HR., Altevogt, BM., editors. Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem. Washington (DC): 2006.
2. Duffy JF, Zitting KM, Czeisler CA. The Case for Addressing Operator Fatigue. Review of human factors and ergonomics. 2015; 10:29–78. [PubMed: 26056516]
3. Goel N, Rao H, Durmer JS, Dinges DF. Neurocognitive consequences of sleep deprivation. Seminars in neurology. 2009; 29:320–339. [PubMed: 19742409]
4. Lyznicki JM, Doege TC, Davis RM, Williams MA. Sleepiness, driving, and motor vehicle crashes. Council on Scientific Affairs, American Medical Association. Jama. 1998; 279:1908–1913. [PubMed: 9634264]
5. Borbely AA. A two process model of sleep regulation. Hum Neurobiol. 1982; 1:195–204. [PubMed: 7185792]
6. Chee MW, et al. Lapsing during sleep deprivation is associated with distributed changes in brain activation. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2008; 28:5519–5528. [PubMed: 18495886]
7. Drummond SP, et al. The neural basis of the psychomotor vigilance task. Sleep. 2005; 28:1059–1068. [PubMed: 16268374]
8. Drummond SP, et al. Sleep deprivation-induced reduction in cortical functional response to serial subtraction. Neuroreport. 1999; 10:3745–3748. [PubMed: 10716202]

9. Padilla ML, Wood RA, Hale LA, Knight RT. Lapses in a prefrontal-extrastriate preparatory attention network predict mistakes. Journal of cognitive neuroscience. 2006; 18:1477–1487. [PubMed: 16989549]
10. Portas CM, et al. A specific role for the thalamus in mediating the interaction of attention and arousal in humans. The Journal of neuroscience : the official journal of the Society for Neuroscience. 1998; 18:8979–8989. [PubMed: 9787003]
11. Thomas M, et al. Neural basis of alertness and cognitive performance impairments during sleepiness. I. Effects of 24 h of sleep deprivation on waking human regional brain activity. Journal of sleep research. 2000; 9:335–352. [PubMed: 11123521]
12. Tomasi D, et al. Impairment of attentional networks after 1 night of sleep deprivation. Cerebral cortex. 2009; 19:233–240. [PubMed: 18483003]
13. Weissman DH, Roberts KC, Visscher KM, Woldorff MG. The neural bases of momentary lapses in attention. Nat Neurosci. 2006; 9:971–978. [PubMed: 16767087]
14. Wu JC, et al. Frontal lobe metabolic decreases with sleep deprivation not totally reversed by recovery sleep. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology. 2006; 31:2783–2792. [PubMed: 16880772]
15. Van Dongen HP, Dinges DF. Investigating the interaction between the homeostatic and circadian processes of sleep-wake regulation for the prediction of waking neurobehavioural performance. Journal of sleep research. 2003; 12:181–187. [PubMed: 12941057]
16. Bernardi G, et al. Neural and behavioral correlates of extended training during sleep deprivation in humans: evidence for local, task-specific effects. J Neurosci. 2015; 35:4487–4500. [PubMed: 25788668]
17. Hung CS, et al. Local experience-dependent changes in the wake EEG after prolonged wakefulness. Sleep. 2013; 36:59–72. [PubMed: 23288972]
18. Pigarev IN, Nothdurft HC, Kastner S. Evidence for asynchronous development of sleep in cortical areas. Neuroreport. 1997; 8:2557–2560. [PubMed: 9261826]
19. Vyazovskiy VV, et al. Local sleep in awake rats. Nature. 2011; 472:443–447. [PubMed: 21525926]
20. Basner M, Dinges DF. Maximizing sensitivity of the psychomotor vigilance test (PVT) to sleep loss. Sleep. 2011; 34:581–591. [PubMed: 21532951]
21. Basner M, Rao H, Goel N, Dinges DF. Sleep deprivation and neurobehavioral dynamics. Current opinion in neurobiology. 2013; 23:854–863. [PubMed: 23523374]
22. Doran SM, Van Dongen HP, Dinges DF. Sustained attention performance during sleep deprivation: evidence of state instability. Archives italiennes de biologie. 2001; 139:253–267. [PubMed: 11330205]
23. Lim J, Dinges DF. Sleep deprivation and vigilant attention. Annals of the New York Academy of Sciences. 2008; 1129:305–322. [PubMed: 18591490]
24. Ma N, Dinges DF, Basner M, Rao H. How acute total sleep loss affects the attending brain: a meta-analysis of neuroimaging studies. Sleep. 2015; 38:233–240. [PubMed: 25409102]
25. Van Dongen HP, Maislin G, Mullington JM, Dinges DF. The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. Sleep. 2003; 26:117–126. [PubMed: 12683469]
26. Andrillon T, Nir Y, Cirelli C, Tononi G, Fried I. Single-neuron activity and eye movements during human REM sleep and awake vision. Nature communications. 2015; 6:7884.
27. Mormann F, et al. Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2008; 28:8865–8872. [PubMed: 18768680]
28. Nir Y, et al. Coupling between Neuronal Firing Rate, Gamma LFP, and BOLD fMRI Is Related to Interneuronal Correlations. Curr Biol. 2007; 17:1275–1285. [PubMed: 17686438]
29. Edwards E, et al. Comparison of time-frequency responses and the event-related potential to auditory speech stimuli in human cortex. J Neurophysiol. 2009; 102:377–386. [PubMed: 19439673]
30. Fisch L, et al. Neural “ignition”: enhanced activation linked to perceptual awareness in human ventral stream visual cortex. Neuron. 2009; 64:562–574. [PubMed: 19945397]

31. Pfurtscheller G, Aranibar A. Event-related cortical desynchronization detected by power measurements of scalp EEG. Electroencephalogr Clin Neurophysiol. 1977; 42:817–826. [PubMed: 67933]
32. Finelli LA, Baumann H, Borbely AA, Achermann P. Dual electroencephalogram markers of human sleep homeostasis: correlation between theta activity in waking and slow-wave activity in sleep. Neuroscience. 2000; 101:523–529. [PubMed: 11113301]
33. Desimone R, Duncan J. Neural mechanisms of selective visual attention. Annual review of neuroscience. 1995; 18:193–222.
34. Steriade M, Timofeev I, Grenier F. Natural waking and sleep states: a view from inside neocortical neurons. Journal of neurophysiology. 2001; 85:1969–1985. [PubMed: 11353014]
35. Nir Y, et al. Regional slow waves and spindles in human sleep. Neuron. 2011; 70:153–169. [PubMed: 21482364]
36. Suthana N, Fried I. Percepts to recollections: insights from single neuron recordings in the human brain. Trends in cognitive sciences. 2012; 16:427–436. [PubMed: 22795560]
37. Harris KD, Thiele A. Cortical state and attention. Nature reviews Neuroscience. 2011; 12:509–523. [PubMed: 21829219]
38. Joshi S, Li Y, Kalwani RM, Gold JI. Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. Neuron. 2016; 89:221–234. [PubMed: 26711118]
39. Wilhelm B, et al. Daytime variations in central nervous system activation measured by a pupillographic sleepiness test. Journal of sleep research. 2001; 10:1–7. [PubMed: 11285049]
40. Parikh V, Kozak R, Martinez V, Sarter M. Prefrontal acetylcholine release controls cue detection on multiple timescales. Neuron. 2007; 56:141–154. [PubMed: 17920021]
41. Fried I, et al. Cerebral microdialysis combined with single neuron and electroencephalographic recording in neurosurgical patients. J Neurosurg. 1999; 91:697–705. [PubMed: 10507396]
42. Suthana NA, et al. Specific responses of human hippocampal neurons are associated with better memory. Proceedings of the National Academy of Sciences of the United States of America. 2015; 112:10503–10508. [PubMed: 26240357]
43. Quiroga RQ, Nadasdy Z, Ben-Shaul Y. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. Neural Comput. 2004; 16:1661–1687. [PubMed: 15228749]
44. Quiroga RQ, Reddy L, Kreiman G, Koch C, Fried I. Invariant visual representation by single neurons in the human brain. Nature. 2005; 435:1102–1107. [PubMed: 15973409]
45. Dawson MW. Fitting the ex-Gaussian equation to reaction time distributions. Behavior Research Methods, Instruments, & Computers. 1988; 20:54–57.
46. Lacouture Y, Cousineau D. How to use MATLAB to fit the ex-Gaussian and other probability functions to a distribution of response times. Tutor Quant Methods Psychol. 2008; 4:35–45.
47. de Curtis M, Avanzini G. Interictal spikes in focal epileptogenesis. Progress in neurobiology. 2001; 63:541–567. [PubMed: 11164621]
48. Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods. 2004; 134:9–21. [PubMed: 15102499]

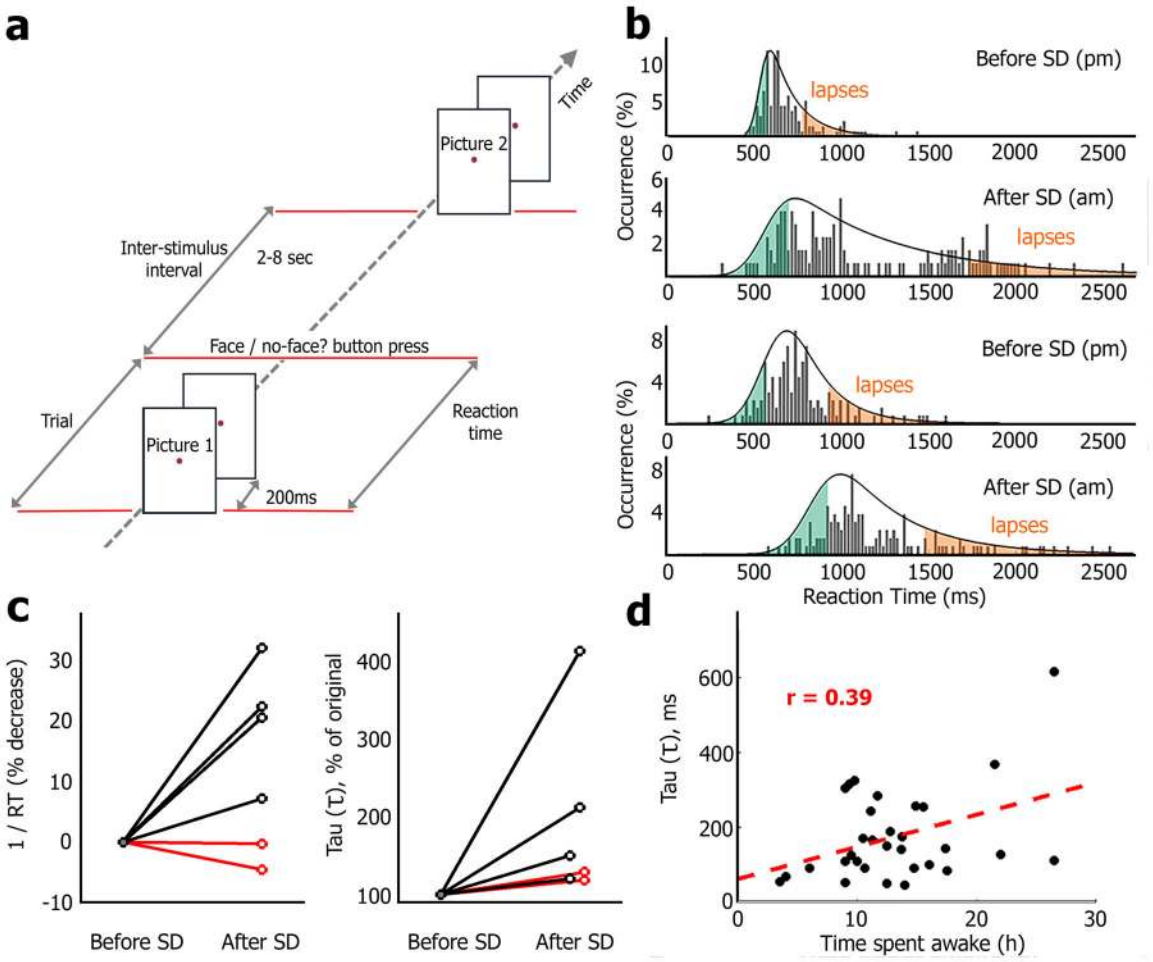

Figure 1. Sleep deprivation (SD) leads to cognitive lapses in a face/non-face categorization Psychomotor Vigilance Task (PVT)

(a) Schematic illustration of the modified PVT where images of people, landmarks, and animals were presented infrequently as participants performed a face/non-face categorization task. **(b)** Distribution of reaction times (RTs) before and after full-night sleep deprivation (SD) in two representative individuals. For each session, an Ex-Gaussian fit (Methods) defines right exponential tail of cognitive lapses (orange, highest RTs) and an equal number of trials with fastest RTs (green) are used for subsequent comparison of neuronal data. **(c)** Left, mean 1/RT (% decrease) in 4 session pairs conducted before/after full-night SD (black) and two sessions conducted before/after normal sleep (red). Note that SD was associated with an increase in mean RT, while normal sleep improved behavioral performance. Right, the tau (τ) parameter (exponential tail in ExGaussian defining cognitive lapses) before/after SD (black, n=4 pairs) and before/after normal sleep (red, n=2 pairs). **(d)** Scatter plot showing significant correlation (Pearson's r=0.39, p<0.03) between tau (τ, ordinate) and time spent awake (abscissa) across all sessions (n=31, not only those conducted before/after SD).

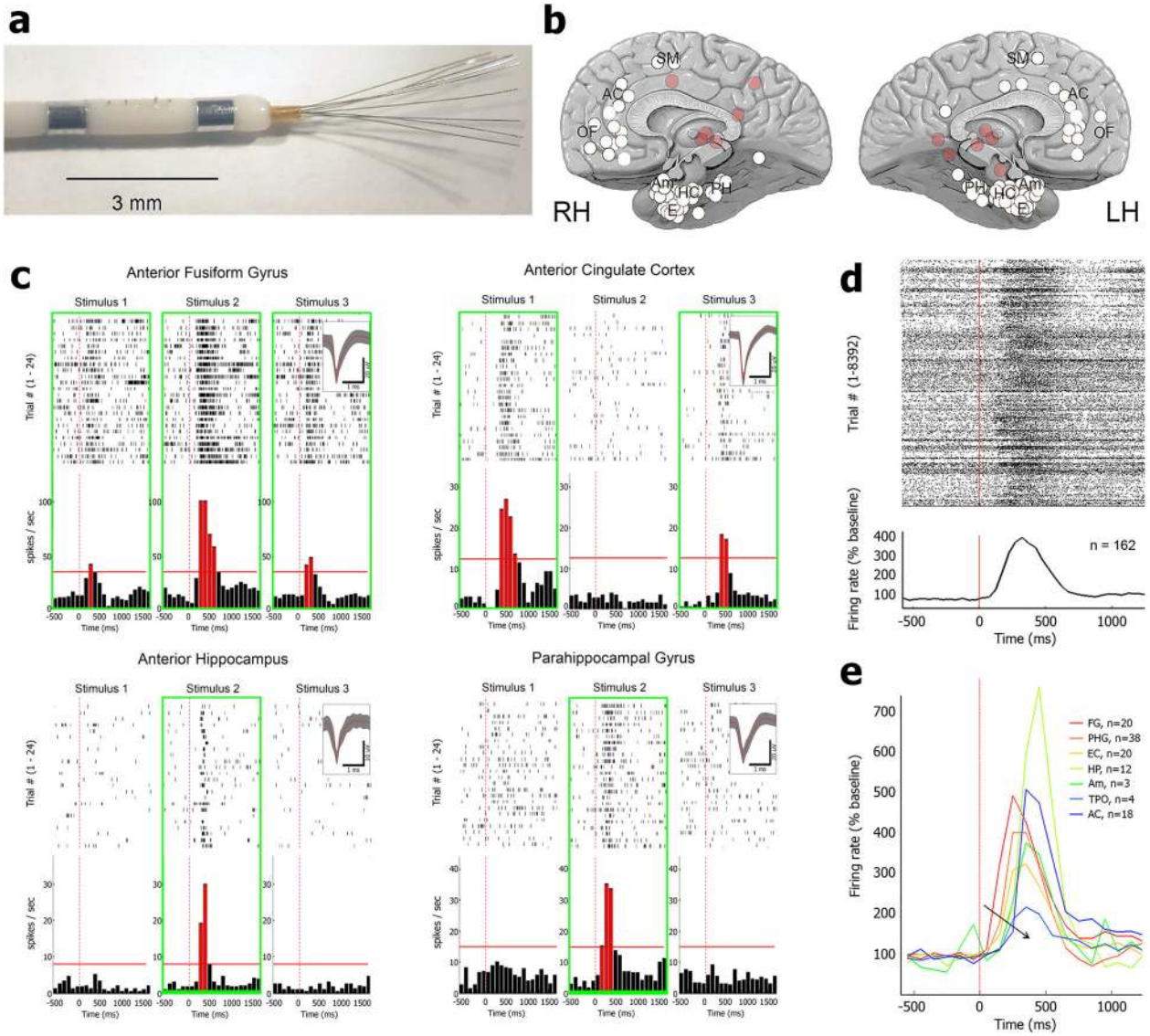

Figure 2. Human single-neuron responses during the face/non-face categorization PVT experiment

(a) Macro-micro depth electrodes with eight 40-μm platinum-iridium microwires protruding 4–5mm from the most distal macro-electrode contact. 6–12 such electrodes were implanted in each patient to simultaneously monitor activity in multiple brain regions **(b)** Overview of 104 depth electrode locations in 12 individuals as seen from medial view. Abbreviations: OF, orbitofrontal cortex; AC, anterior cingulate; SM, supplementary motor; PH, parahippocampal gyrus; HC, hippocampus; E, entorhinal cortex; Am, amygdala; LH, left hemisphere; RH, right hemisphere. Opaque red circles mark more lateral regions such as superior temporal gyrus. **(c)** Four representative examples (raster plots and peri-stimulus-time histograms, PSTH) of single-unit spiking responses to pictures recorded from the Anterior Fusiform Gyrus (top left), Anterior Hippocampus (bottom left), Anterior Cingulate Cortex (top right), and Parahippocampal Gyrus (bottom right). Green boxes mark stimuli eliciting significant responses (red bars) above baseline firing (horizontal red lines) while

insets show action potential waveforms. **(d)** Average response (raster plot and PSTH) across all neurons (n=162) tagged as 'responsive', for pictures that were effective in driving a response, reveals a robust increase in spike discharges around 200–500ms after stimulus onset. **(e)** Average response, shown separately for each brain region monitored, reveals an orderly progression of temporal latencies (black arrow, hot-to-cold colors) from high-order visual cortex to hippocampus and frontal lobe (FG, anterior Fusiform Gyrus; PHG, Parahippocampal Gyrus; EC, entorhinal cortex; HP, hippocampus; Am, amygdala; TPO, temporal-parietal-occipital junction; AC, anterior cingulate cortex).

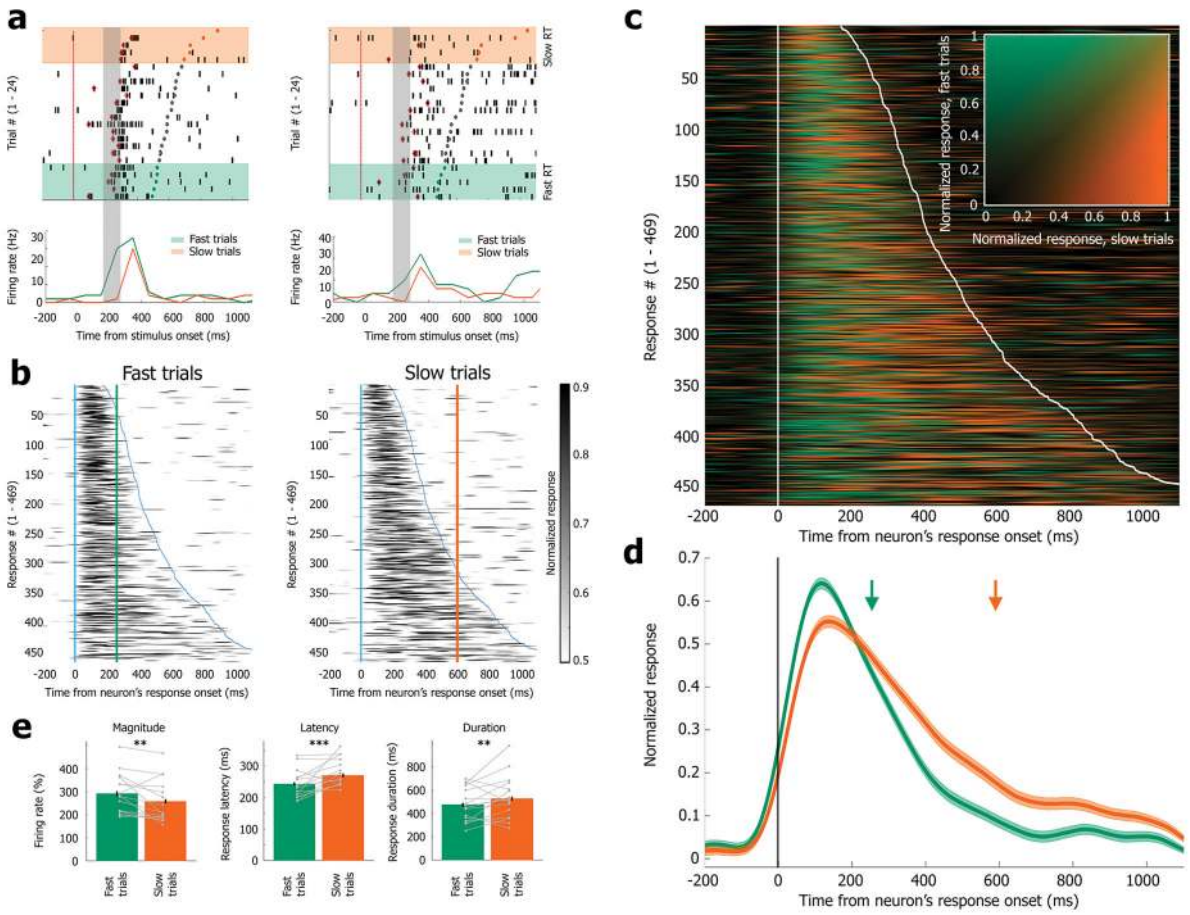

Figure 3. Reduced, delayed, and lengthened single-unit responses during cognitive lapses

(a) Spiking responses (raster and PSTH) in fast trials (lowest RTs, green) vs. slow trials (highest RTs, orange) of two representative neurons in the anterior hippocampus and the parahippocampal gyrus (same neurons as bottom rows in Figure 2C). Trials in raster plot are sorted based on RTs in each trial (slowest on top). Black ticks, action potentials; Open red circles, response latency detected automatically; Green/gray/orange circles, behavioral response in fast/other/slow trials, respectively. Orange rectangle shading, slow trials; Green rectangle shading, fast trials. **(b)** Normalized PSTH of all responses (each row represents a response to one of 469 stimuli; 162 responsive neurons) during fast trials (left) and slow trials (right). Responses are aligned to each neuron's response onset across all trials (x-axis), and amplitude (color scale) is normalized to each neuron's peak response to go beyond variability across neurons in response timing and amplitude. Vertical and diagonal blue lines mark average time of response onset, and response termination, respectively, for each neuron (sorted by response duration). Green and orange vertical lines mark mean behavioral RT in fast and slow trials, respectively. **(c)** Color superposition of PSTH responses (each row represents a response to one of 469 stimuli; 162 responsive neurons) in fast and slow trials. Responses are aligned (x-axis) and normalized (y-axis) as in (b). Color brightness (inset legend) represents firing rate magnitude, while hue (green vs. orange) represents stronger responses during fast vs. slow trials at that time (Methods). Vertical and diagonal white lines mark average time of response onset and response termination, respectively, for each neuron

(sorted by response duration). Note stronger earlier response in fast trials (green dominance around onset) vs. delayed and lengthened response in slow trials (orange dominating later). **(d)** Grand-mean PSTH of all responses (n=469 responses in 162 neurons) in fast trials (green) and slow trials (orange). Responses are aligned (x-axis) and normalized (y-axis) as in (b). Green and orange arrows mark mean behavioral RT in fast and slow trials, respectively. **(e)** Quantification of response magnitude (left), response latency (middle), and response duration (right) in individual responsive neurons during fast trials vs. slow trials (N=376 pictures in 142 units). Different N values with respect to previous panels stem from single-trial analysis (Methods). Slow trials are associated with statistically significant firing rate reduction (**, $p<0.005$, Wilcoxon signed-rank test), increased temporal latency (***, $p<0.0005$, Wilcoxon signed-rank test), and longer response duration (**, $p<0.005$, Wilcoxon signed-rank test). Gray dots/lines depict 16 individual sessions with at least 5 unit responses.

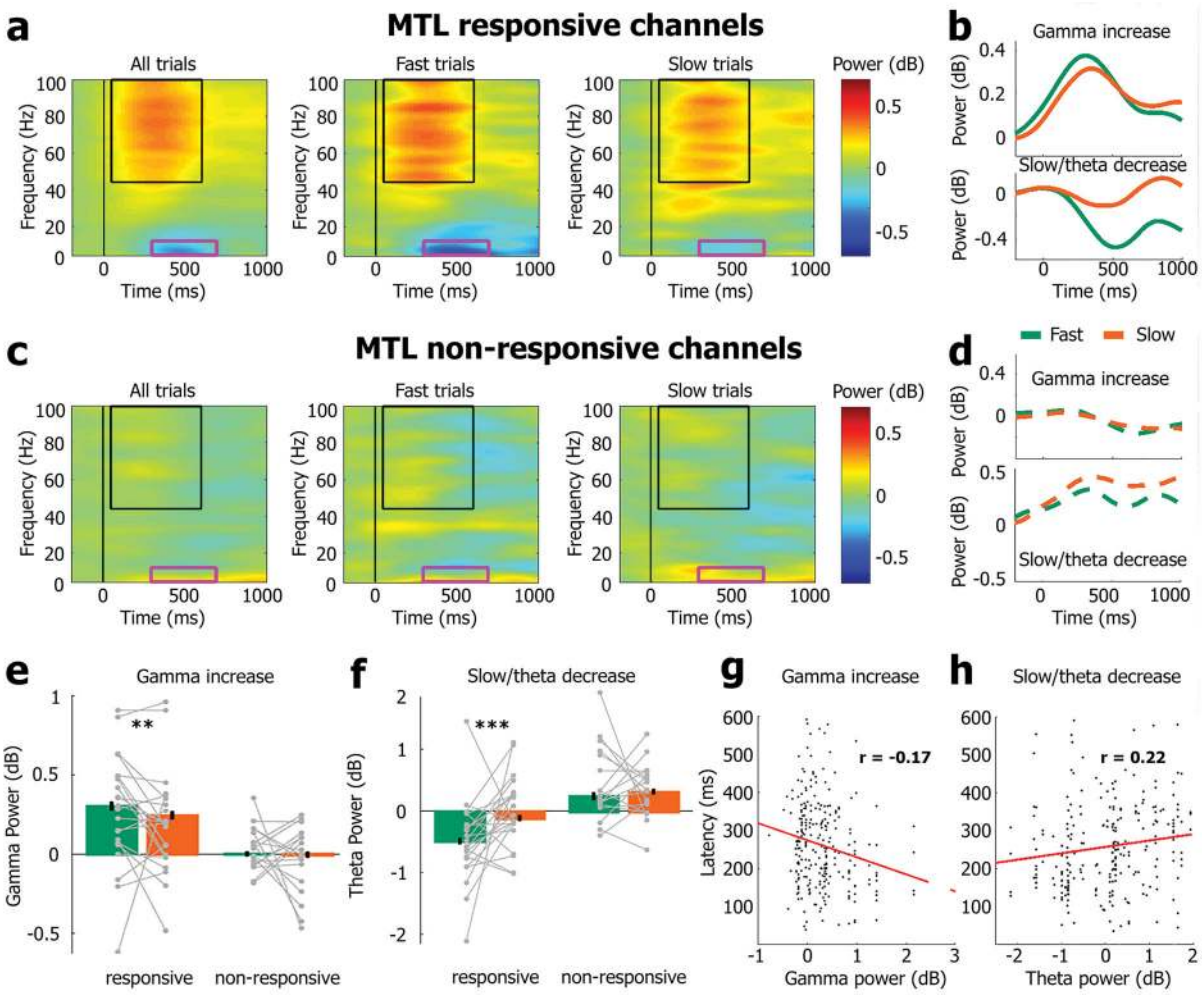

Figure 4. Cognitive lapses are associated with weaker gamma power increase and weaker slow/theta power decrease in MTL LFPs

(a) Time-frequency decomposition of induced power changes in local field potentials (LFPs) of MTL responsive channels (n=270 channels in 31 sessions). Columns denote the average power changes for all trials (left), fast trials (lowest RTs, middle), and slow trials (highest RTs, right). In each subpanel, hot and cold colors mark increases and decreases in power, respectively. Black rectangles mark stimulus-induced increase in gamma frequency (>45Hz) power around 50–600ms after stimulus onset. Pink rectangles mark stimulus-induced decreased power in the slow/theta frequency range (2–10Hz) around 300–700ms after stimulus onset. Slow trials are associated with weaker gamma power increase and weaker slow/theta power decrease. **(b)** Time course of gamma power increase (top) and slow/theta power decrease (bottom) for fast trials (green) vs. slow trials (orange). **(c)** Same as (a) for neighboring MTL non-responsive channels (n=198 channels in 31 sessions). Note the absence of significant power modulations during the same slow trials in neighboring channels. **(d)** Same as (b) for neighboring MTL non-responsive channels. **(e)** Quantification (median) of gamma power increases (45–100Hz; 50–600ms) for responsive (left) and non-responsive MTL channels (right). Asterisks show significant differences (Wilcoxon signed-

rank tests comparing fast trials with slow trials; **, $p<0.007$). **(f)** Quantification (median) of slow/theta power decrease (2–10Hz; 300–700ms) for responsive (left) and non-responsive MTL channels (right). In panels (e) and (f), error bars denote SEM computed across LFP channels (n=270 and 198 for responsive and non-responsive channels, respectively), and gray dots/lines mark 22 individual sessions (responsive channels) and 17 individual sessions (unresponsive channels) that had at least 5 LFP channels. Asterisks show significant differences (Wilcoxon signed-rank tests comparing fast trials with slow trials; ***, $p<10^{-7}$). **(g)** Scatter plot of single-neuron response latency (y-axis) vs. strength of gamma power increase (x-axis) reveals that during slow trials, increased latency in spiking responses is significantly correlated with weaker increase in LFP gamma (Spearman coefficient r=−0.17, p=0.007, n=255 pictures that elicited significant responses across 87 units and 21 sessions, see Methods). **(h)** Scatter plot of single-neuron response latency (y-axis) vs. strength of slow/theta power decrease (x- axis) reveals that during slow trials, increased latency in spiking responses is significantly correlated with increased slow/theta LFP power (Spearman coefficient: r=0.22, $p=4.5*10^{-4}$ in n=255 pictures that elicited significant responses across 87 units and 21 sessions).