nature human behaviour

ARTICLES

https://doi.org/10.1038/s41562-022-01335-5

Theta oscillations shift towards optimal frequency for cognitive control

Mehdi Senoussi[1]✉, Pieter Verbeke[1], Kobe Desender[1,2,3], Esther De Loof[1], Durk Talsma[1] and Tom Verguts[1]

Cognitive control allows to flexibly guide behaviour in a complex and ever-changing environment. It is supported by theta band (4–7 Hz) neural oscillations that coordinate distant neural populations. However, little is known about the precise neural mechanisms permitting such flexible control. Most research has focused on theta amplitude, showing that it increases when control is needed, but a second essential aspect of theta oscillations, their peak frequency, has mostly been overlooked. Here, using computational modelling and behavioural and electrophysiological recordings, in three independent datasets, we show that theta oscillations adaptively shift towards optimal frequency depending on task demands. We provide evidence that theta frequency balances reliable set-up of task representation and gating of task-relevant sensory and motor information and that this frequency shift predicts behavioural performance. Our study presents a mechanism supporting flexible control and calls for a reevaluation of the mechanistic role of theta oscillations in adaptive behaviour.

Cognitive control permits adapting behaviour to task demands, crucial in an ever-changing environment. The flexibility of such a fundamental cognitive ability lies at the core of intelligent behaviour. Cognitive control is supported by neural oscillations in the theta band (4–7 Hz)[1,2] (also called frontal midline theta) that coordinate distant neural populations to create task-relevant functional networks through synchronization[3–7]. Medial frontal cortex (MFC) generates theta oscillations when cognitive control is needed, that is, during conflict or in preparation of a difficult task[1]. Task rules and goals dictating behaviour are instead encoded in lateral frontal cortex (LFC)[8–10]. The coordination of these two areas through theta-rhythmic processes has been shown to support successful task performance[11]. It has been proposed that task-relevant functional networks are established through top-down gating from these frontal areas by synchronizing distant neural populations, allowing efficient communication, that is, communication through coherence (CTC)[3,12].

Theta oscillations thus play a critical role in the implementation of cognitive control, but to be adaptive, theta oscillation characteristics must change with task demands. However, the exact neural mechanisms that support flexible control remain largely unknown. Most research has focused on theta amplitude, showing that it increases after conflicts and errors, causing subsequent neural adaptation leading to better task performance[1].

Critically, a second essential aspect of theta oscillations, their peak frequency within the 4–7 Hz range, has occasionally been reported to vary across tasks and participants[13–15]. However, most studies report band-average theta power per condition, which precludes observing changes in peak theta frequency across conditions. Moreover, estimating shifts in peak frequency from conventional representation of spectral data (for example, power spectra or time–frequency maps) is nontrivial and must avoid confounding factors such as changes in the aperiodic component of the power spectra[16]. It therefore remains unclear whether reliable theta peak differences exist. Finally, this variability and its mechanistic consequences are commonly ignored, and no theoretical account has considered its role in cognitive control. To address this gap, we draw from two prominent frameworks: biased competition (BC)[17] and CTC[12]. We build a computational model where theta oscillations orchestrate competition between task representations, which in turn guides CTC to set up task-relevant functional networks. Model simulations show that, depending on task demands, different theta frequencies are optimal for task performance. We tested model predictions on behavioural and electrophysiological data and confirmed that the frequency of theta oscillations shifts adaptively towards optimal frequency depending on task demands.

Results

Theta frequency controls reliable task implementation. We designed a stimulus–action mapping task (Fig. 1a) wherein on each trial, a different mapping (that is, a rule, with variable difficulty) must be established. The task consists of reporting the tilt of one of two gratings, clockwise (CW) or counter-CW (CCW) from the vertical axis, using the index or middle finger of one of both hands. On each trial, a two-letter cue instructed the rule, that is, which was the target grating (left (L) or right (R), top letter) and which hand to use (L or R, bottom letter). We thus manipulated task difficulty: same-side cues (that is, RR and LL, indicating top and bottom letter, respectively) were easier than different-side cues (LR and RL).

Our model consists of five units (Fig. 1b): two control units (LFC and MFC), two processing units (sensory and action) and an integrator unit. In LFC, cues activate instruction nodes, which themselves activate rule nodes. Rule nodes form a competitive accumulator network[18] that implements BC: In a Stroop-like manner, the connectivity between instruction nodes induces stronger competition between rule nodes for different-side than same-side rules. Importantly, rule node competition is orchestrated by theta oscillations generated by the MFC unit: competition is (re)initiated when MFC theta exceeds a processing threshold (Fig. 2a,b). Each rule node points to rule-relevant processing modules. Processing nodes oscillate at gamma frequency. Rule nodes gate communication between sensory and action units through CTC[12,19], thereby

[1]Department of Experimental Psychology, Ghent University, Ghent, Belgium. [2]Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. [3]Brain and Cognition, KU Leuven, Leuven, Belgium. ✉e-mail: mehdi.senoussi@ugent.be

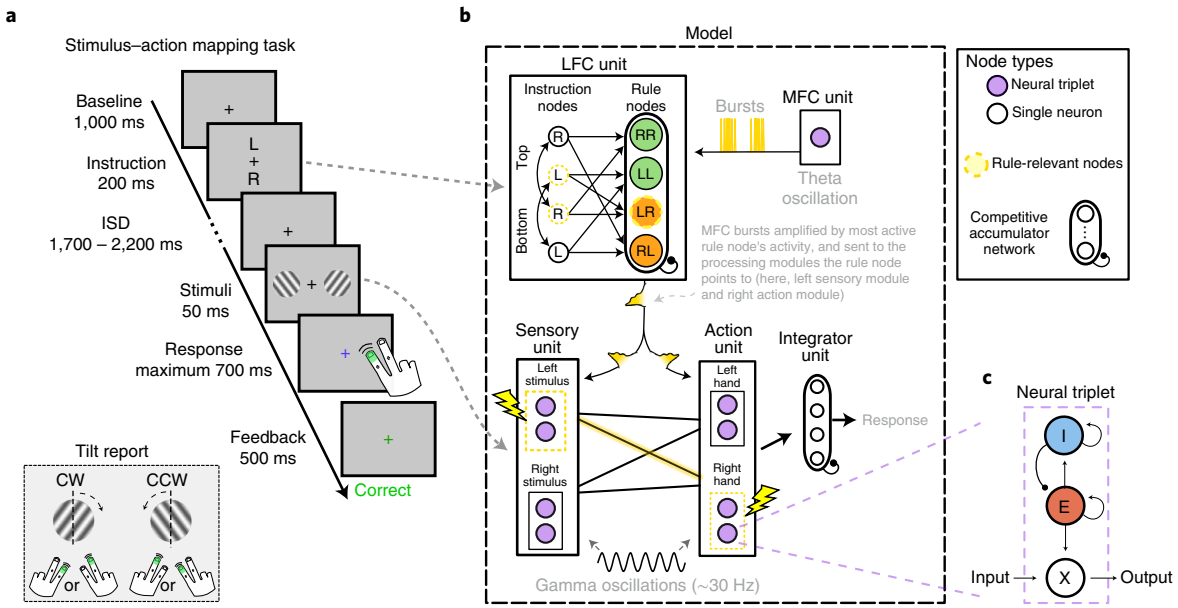


Fig. 1 | Task and model structure. a, The stimulus–action mapping task. Each trial starts with a cue instructing the mapping to use. In this example, the rule is LR, instructing to report the left grating's tilt with the right hand. **b**, Model architecture. **c**, Each node of MFC and processing (sensory and action) units is a neural triplet composed of one excitatory (E), one inhibitory (I) and one rate neuron (x). The E–I pair generates oscillations (whose frequency depends on their coupling parameter). MFC bursts are sent to E neurons. Rate neurons receive input from, and send output to, other nodes' rate neurons. The activity (output) of a rate neuron is modulated by its E neuron.

implementing the instructed mapping, by means of phase-resetting bursts emitted by MFC at theta oscillation peaks[3,19] (Fig. 2b–d). The integrator unit constitutes a competitive accumulator network[18] that accumulates information received from action nodes and triggers a response once one of the integrator nodes reaches a threshold (Fig. 2e).

Crucially, with a fast theta frequency, for example, 7 Hz, rule nodes gate processing modules frequently, shortening 'off' periods in which rule-relevant processing nodes desynchronize, at the cost of shorter competition windows. With a slow theta frequency, for example, 4 Hz, gating is imposed less frequently but competition windows are longer. Due to BC, one rule will win the competition. However, for difficult rules, resolving the competition will take more time, that is, require longer competition windows. In our task, different-side rules are more difficult, so the model achieves better performance at slower theta frequencies where competition is long enough for the correct rule node to win (Extended Data Fig. 1a,b). In contrast, for easy rules, competition is won quickly, thus higher theta frequencies yield better performance as rule-relevant nodes are frequently gated, reducing 'off' periods. Hence, an optimal agent would shift theta frequency depending on task demands.

Model simulations (Fig. 3a) confirmed that, for difficult rules, the model achieves optimal accuracy at a slow theta frequency, whereas for easy rules, a fast theta is optimal ($W = 105.5$, $P < 0.001$, $r = 0.64$, 95% CI 1.00–2.00; Fig. 3b). Fits from the drift–diffusion model (DDM) on model data (Methods) showed that only drift rate exhibited this theta frequency–rule difficulty interaction (Extended Data Fig. 2a), refuting a speed–accuracy trade-off explanation. Theta amplitude alone could not explain this result as theta amplitude only negligibly affected competition window length relative to frequency (Extended Data Fig. 1c,d).

Furthermore, theta-rhythmic gating of processing nodes should yield better model performance shortly after a burst, that is, at theta oscillation peaks (Fig. 3c). By varying the instruction–stimulus delay (ISD), to sample model performance at different phases of the theta-rhythmic process[20,21], we showed that model accuracy oscillates at a frequency closely matching MFC theta frequency (Fig. 3d and Supplementary Fig. 1).

These simulations lead to two key behavioural and neural predictions. First, oscillations of accuracy-by-ISD should shift towards optimal theta frequency depending on task demands. Second, frontal theta oscillations should also exhibit this effect, and the degree to which theta frequency shifts according to task demands should be predictive of subsequent task performance.

Frequency shift in behavioural performance oscillations. In an experiment on human participants (dataset 1), we first confirmed that rules varied in difficulty (Fig. 4a). There was a significant target location–hand interaction in accuracy (RR and LL easier than LR and RL; $F(1, 33) = 27.82$, $P < 0.001$, $\eta^2 = 0.236$) and a main effect of hand ($F(1, 33) = 4.33$, $P = 0.045$, $\eta^2 = 0.012$). Consistent with model simulations, only drift rate exhibited this interaction (Extended Data Fig. 2b). We therefore used accuracy as our dependent variable. To test model predictions on behavioural oscillations, we computed peak theta frequency of accuracy-by-ISD (Methods and Supplementary Fig. 2a). As predicted, we found a significant target location–hand interaction ($F(1, 33) = 6.51$, $P = 0.015$, $\eta^2 = 0.047$), showing that accuracy oscillated at a slower theta frequency for difficult rules (LR and RL; Fig. 4b), and no main effect.

Frequency shift in frontal theta predicts task performance. Next, we investigated whether neural theta exhibited this frequency shift due to task demands. We extracted electroencephalogram (EEG) theta peak frequency in a 1 s prestimulus window from an electrode cluster exhibiting significantly higher theta power in correct than incorrect trials ($P < 0.001$; Fig. 4c and Methods). As predicted, peak theta frequency in correct trials significantly decreased from same-side to different-side rules ($F(1, 33) = 18.96$, $P < 0.001$, $\eta^2 = 0.107$; Fig. 4d and individual participant spectra in Extended Data Fig. 3). Although peak theta frequency differed numerically between different-side rules (that is, LR and RL), this difference was not statistically significant ($W = 211$, $P = 0.139$, $r = -0.29$,

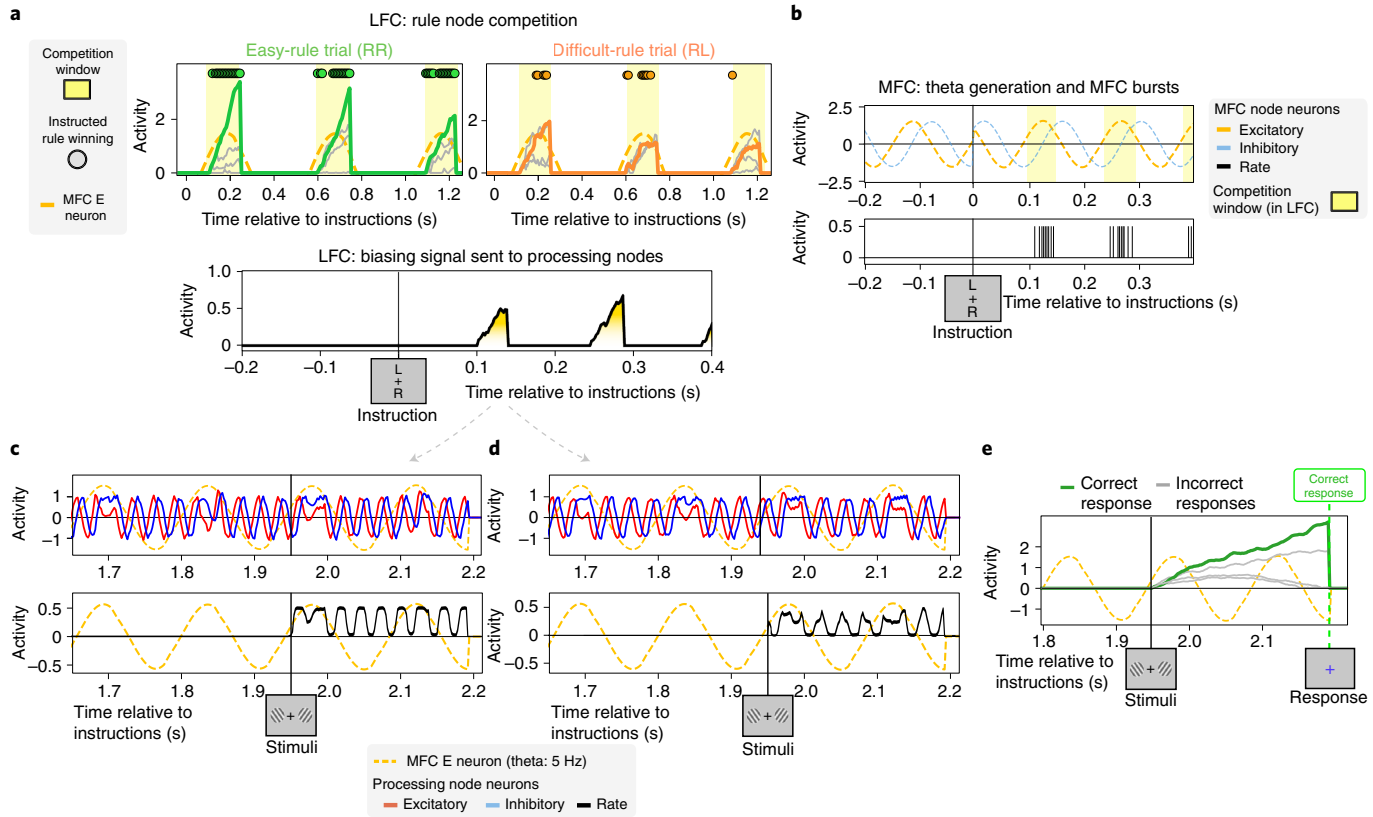


Fig. 2 | Model dynamics. a, The two upper panels represent rhythmic BC in rule nodes at instructions onset (from two different trials): three cycles at a fast MFC theta frequency (7 Hz) to illustrate difference in dynamics between easy and difficult rules. Top left panel (easy rule): the rule node corresponding to the instructed rule in this trial (green curve) rapidly wins the competition over other rule nodes (grey curves). Top right panel (difficult rule): the rule node corresponding to the instructed rule in this trial (orange curve) struggles to win the competition and often loses to other rule nodes (grey curves). Yellow areas represent competition window opened by MFC theta oscillations. Dashed yellow line represents MFC E neuron activity (see **b**). Each dot above curves represents a time point at which the instructed rule's node was winning the competition. The bottom panel represents the biasing signal sent to the processing nodes. This corresponds to the burst sent by MFC (bottom plot in **b**) multiplied by the most activated rule node's activity (the MFC theta frequency was set at 5 Hz for this panel). **b**, Activity of the MFC neural triplet at a theta frequency of 5 Hz. Top panel represents activity of the E–I pair. The phase of both neurons is reset at instructions onset. Bottom panel represents activity of the rate neuron. **c**,**d**, Time course of neural triplet activity of the rule-relevant sensory (**c**) and action (**d**) nodes around stimuli presentation with an MFC theta frequency of 5 Hz. The yellow curve representing MFC theta oscillations is for illustration only, its activity is scaled to fit within these plots. **e**, Time course of the four integrator nodes, representing each of the four possible responses to the task, around stimuli presentation. The green line represents the correct response in the simulated trial.

95% CI −0.72 to −0.14). Furthermore, contrasting correct and incorrect trials revealed that higher theta frequency improved performance in same-side rules whereas a lower theta frequency improved performances in different-side rules ($F(1, 33) = 4.62$, $P = 0.039$, $\eta^2 = 0.036$; Fig. 4d). Finally, across participants, the degree to which theta frequency shifted from difficult to easy rules positively correlated with overall accuracy ($r(32) = 0.49$, $P = 0.004$, 95% CI 0.17–0.71; Fig. 4e), indicating that a higher sensitivity of theta frequency to rule difficulty improved task performance. These analyses were carried out by using the FOOOF toolbox[16] to estimate peak and power of theta oscillations. Additional control analyses revealed that our results were robust and observable without using this toolbox (that is, by estimating theta peak frequency on raw spectra; Methods and Supplementary Fig. 7). These results cannot be explained by changes in theta power alone as both peak and power were estimated independently over the $1/f$ spectrum, where f is frequency (Supplementary Fig. 2b and Control analyses section).

Theta frequency shift generalizes to other tasks. Having established a robust effect of task demands on theta frequency in our stimulus–action mapping task, we tested the generality of this mechanism, namely a decrease of theta frequency for difficult tasks, to other cognitive control tasks. First, we reanalysed previously published data[22] from an experiment in which 17 participants performed an arithmetic task, preceded by a cue indicating whether the arithmetic operation was going to be easy or difficult (dataset 2; Fig. 5a). Different from our original experiment (that is, dataset 1), in dataset 2, only two levels of difficulty were used, thus allowing us to test whether theta frequency is lower following a difficult compared with an easy cue. There was a significant effect of difficulty on error rates and on reaction times (see the original article[22] for details). To use a comparable time window for the analysis of the EEG data (relative to dataset 1), we selected a 1 s segment of EEG data in the post-cue interval. This segment was centred around the time point at which the difference in theta power between difficult and easy conditions was the highest (2,000 ms post-cue onset; fig. 4a in ref. [22]). Thus, we considered EEG data in the 1,500–2,500 ms segment post cue onset. Furthermore, because of the low number of incorrect responses (error rates of 1% and 6% for the easy and difficult conditions, respectively), we decided not to use the correct–incorrect contrast as in dataset 1 (Fig. 4c) and chose an a priori electrode (FCz) on the basis of prior findings in theta oscillations in cognitive control[1,23–25]. Due to the absence of identified theta oscillations using the FOOOF toolbox in the easy condition of

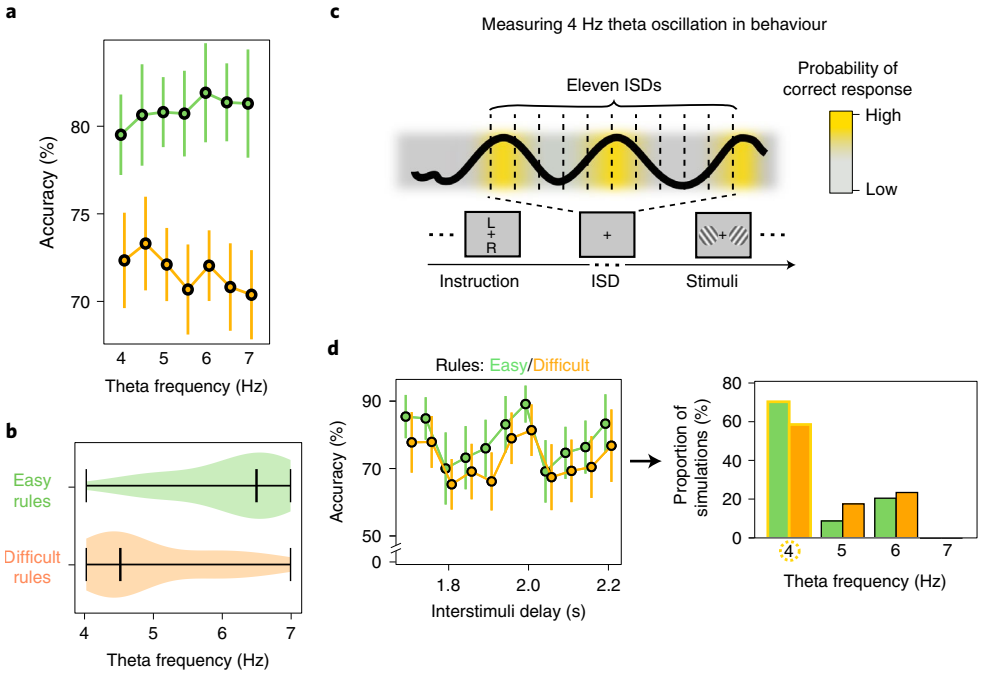


Fig. 3 | Model simulations. a, Model accuracy by rule difficulty across theta frequencies. Data presented as mean values; error bars represent s.d. computed over $n = 34$ simulations per frequency. Green curve represents easy (same-side) rules; orange curve represents difficult (different-side) rules. **b**, Violin plots representing optimal theta frequency, that is, yielding highest model accuracy, per rule difficulty. For difficult rules, the model achieves optimal accuracy at a slow theta frequency, whereas for easy rules, a fast theta is optimal ($n = 34$ simulations per theta frequency, two-sided Wilcoxon signed-rank test: $W = 105.5$, $P < 0.001$, $r = 0.64$, 95% CI 1.00–2.00). Data presented as violin plots, left- and rightmost bars represent extrema and middle bar represents the median. Distribution density is represented by violin plot width. **c**, Measuring theta oscillations in behaviour through densely distributed ISD. **d**, Accuracy-by-ISD by rule difficulty (left) and estimated peak frequency (right). Data presented as mean values. Error bars represent standard deviation computed over $n = 34$ simulations. The yellow highlighting in the right panel around bars and 4 Hz represent the MFC frequency in the model for these simulations.

3 participants, we analysed 14 participants in total in dataset 2. Confirming the model predictions and the observations from dataset 1, we found a lower peak theta frequency in correct trials in the difficult compared with easy condition ($W = 86$, $P = 0.017$, $r = 0.64$, 95% CI 0.01–0.14; Fig. 5b).

Second, we reanalysed another published dataset[26] from an experiment in which 33 participants performed a go–no-go task, where each trial was preceded by a cue indicating whether the upcoming stimulus was a certain-go (that is, a go stimulus with 100% certainty) or a maybe-go (that is, a no-go stimulus with 25% certainty, dataset 3; Fig. 5c). There was a significant effect of cue type on error rates (see the original article[26] for details). We tested whether theta frequency is lower following a maybe-go cue compared with a certain-go cue (followed by a no-go stimulus). As for datasets 1 and 2, we estimated peak theta frequency in a 1 s segment preceding stimulus onset. Similar to dataset 2, the number of incorrect responses was low (1.7% error rates in the certain-go condition), thus we used electrode FCz. Again, confirming our findings from the model and datasets 1 and 2, we found a higher peak theta frequency following certain-go cues compared with maybe-go cues ($W = 358$, $P = 0.039$, $r = 0.36$, 95% CI −0.01 to 0.08; Fig. 5d).

Control analyses. We showed that peak theta frequency decreases with task difficulty. A recent study demonstrated that, in the case of posterior alpha oscillations, amplitude and frequency are intrinsically related[27] due to the thalamocortical circuits thought to generate alpha oscillations[28]. One concern could be that such a relationship also exists in the case of midfrontal theta oscillations, thereby confounding frequency and power. We thus verified whether peak theta amplitude exhibited the same pattern of decrease with task difficulty. We carried out the same analysis procedure that was used for peak theta frequency on peak theta amplitude (Fig. 6) and showed that there was no statistically significant difference in peak theta amplitude between easy and difficult conditions in any of the three datasets (all P values >0.078, uncorrected for multiple comparisons). This result shows that peak theta amplitude could not account for the decrease in peak theta frequency across conditions.

Additionally, we tested whether the shift in peak theta frequency could be confounded by amplitude or frequency of nearby frequency bands. For both the delta and alpha frequency bands, we followed the same procedure as for the main results on the theta band but instead analysed peaks in the 1–3 Hz range (delta) or 8–12 Hz range (alpha band). No statistically significant decrease with task difficulty was found for the delta or alpha band, neither in peak frequency (all P values >0.091, uncorrected for multiple comparisons; Extended Data Fig. 3a,c,e,g,i,k) nor in peak amplitude (all P values >0.200, uncorrected for multiple comparisons; Extended Data Fig. 3b,d,f,h,j,l).

These control analyses therefore suggest that the shift of peak theta frequency with task difficulty occurs independently of changes in theta amplitude or changes in nearby frequency bands.

Discussion

We have identified an adaptive mechanism allowing flexible cognitive control. We propose a computational model, test its predictions in behavioural and electrophysiological data and show that theta oscillations lawfully adapt to task demands by shifting towards optimal frequency for task performance. Moreover, we replicate this finding in two independent datasets implementing entirely different tasks (arithmetic operations and response inhibition) and show that

the shift of theta frequency according to task demands is a general mechanism involved in a wide range of cognitive processes. Finally, we controlled for possible confounding factors such as amplitude modulations and changes in nearby frequency bands and showed that the shift in theta peak frequency took place independently from other changes in oscillatory activity.

These findings are in line with evidence that frequency of neural oscillations adapts to external demands, for example, perceptual demands in alpha band[29,30], and is related to short-term memory capacity in theta band[31]. Our study complements and extends our understanding of how neural oscillations support cognitive processes by providing a mechanistic account allowing to simulate and test further hypotheses. An exciting avenue for future research lies in characterizing how adaptive shifts in theta frequency relate to cross-frequency coupling dynamics[32], for example, between theta and gamma oscillations.

A related body of work has investigated the role of theta peak frequency in working memory processes. Indeed, theta oscillations originating from medial temporal lobe and basal forebrain structures (for example, hippocampus and septum) have been hypothesized to support the maintenance of ordinal information in an item sequence in working memory[31]. According to this theory, the phase of theta oscillations structures the (re)activation of distinct neural populations oscillating at gamma frequency, each representing an item of the maintained sequence. This theory thus predicts that a slower theta frequency, leading to longer periods in which items could be nested, would increase working memory capacity (see also ref. [33] for a discussion of oscillatory frequency and cognitive resources). Some studies have confirmed this prediction empirically by showing that higher working memory loads led to a reduction of theta frequency[34,35]. Moreover, a recent study causally tested this prediction using transcranial alternating current stimulation (tACS)[36] and showed that stimulating a frontoparietal network at a slow (4 Hz) versus fast (7 Hz) theta frequency led to increase in working memory capacity. Although theta oscillations that support working memory and cognitive control serve different purported roles (that is, structuring maintained information in working memory versus synchronizing for communication in cognitive control) and have distinct neural origins (hippocampus/septum in working memory versus MFC in cognitive control), both

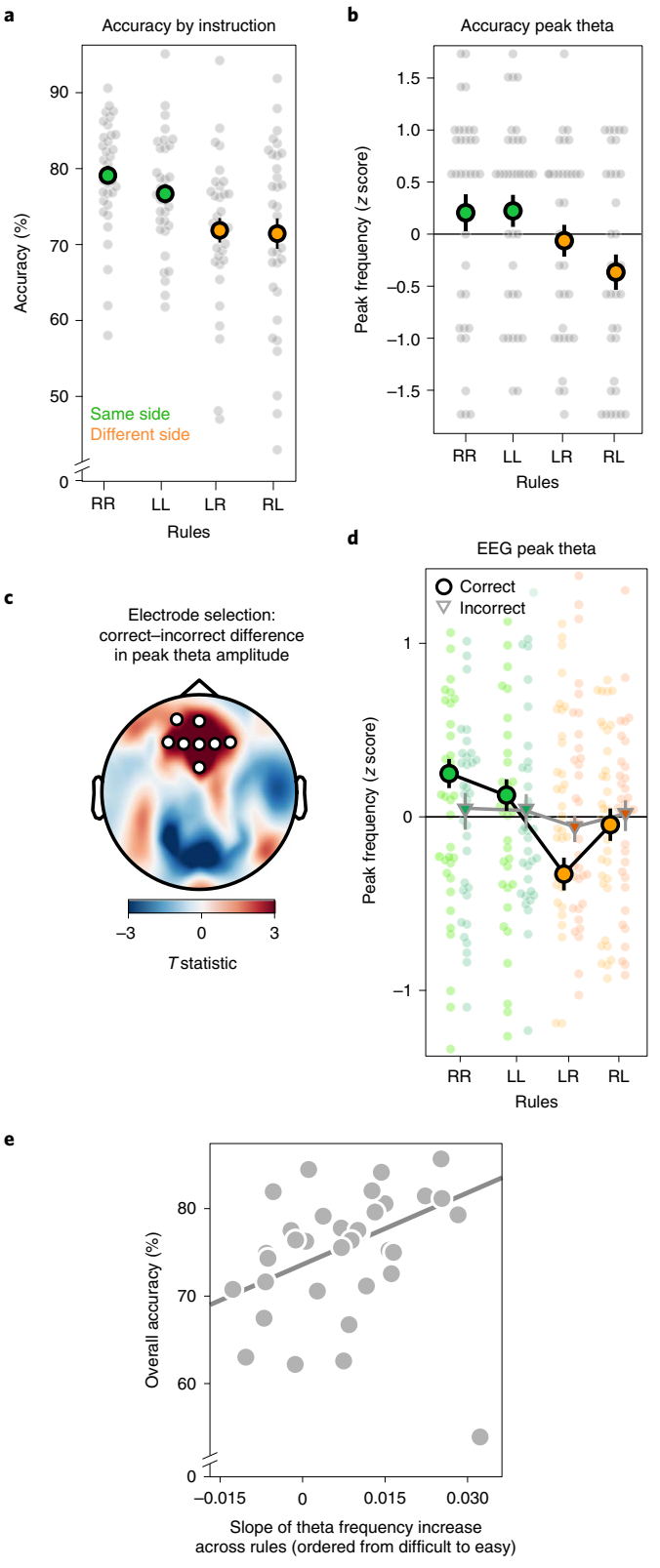


Fig. 4 | Testing model predictions in behaviour and EEG. a, Overall accuracy. There was a significant target location–hand interaction in accuracy (RR and LL easier than LR and RL; two-way repeated-measure ANOVA: $F(1, 33) = 27.82$, $P < 0.001$, $\eta^2 = 0.236$) and a main effect of hand ($F(1, 33) = 4.33$, $P = 0.045$, $\eta^2 = 0.012$). Data presented as mean values. Error bars represent s.e.m. computed over $n = 34$ participants. Smaller grey dots represent individual participants' data. **b**, Peak frequency of oscillations in accuracy-by-ISD: there was a significant target location–hand interaction (two-way repeated-measure ANOVA: $F(1, 33) = 6.51$, $P = 0.015$, $\eta^2 = 0.047$). Data presented as mean values. Error bars represent s.e.m. computed over $n = 34$ participants. Smaller grey dots represent individual participants' data. **c**, Frontal cluster of electrodes with increased theta amplitude: significantly higher theta amplitude in correct than incorrect trials (cluster test, $P < 0.001$). **d**, Theta peak frequency by rule in frontal electrodes cluster. Peak theta frequency in correct trials significantly decreased from same-side to different-side rules (two-way repeated-measure ANOVA: $F(1, 33) = 18.96$, $P < 0.001$, $\eta^2 = 0.107$). Comparing correct and incorrect trials, we found that higher theta frequency improved performance in same-side rules, whereas a lower theta frequency improved performances in different-side rules (two-way repeated-measure ANOVA: $F(1, 33) = 4.62$, $P = 0.039$, $\eta^2 = 0.036$). Data presented as mean values. Error bars represent s.e.m. computed over $n = 34$ participants. Smaller colored dots represent individual participants' data for each rule and for correct and incorrect trials separately. **e**, Correlation between theta peak slope across rule difficulty and overall accuracy: the degree to which theta frequency shifted from difficult to easy rules positively correlated with overall accuracy (robust Spearman correlation: $r(32) = 0.49$, $P = 0.004$, 95% CI 0.17–0.71). Each dot represents a single participant.

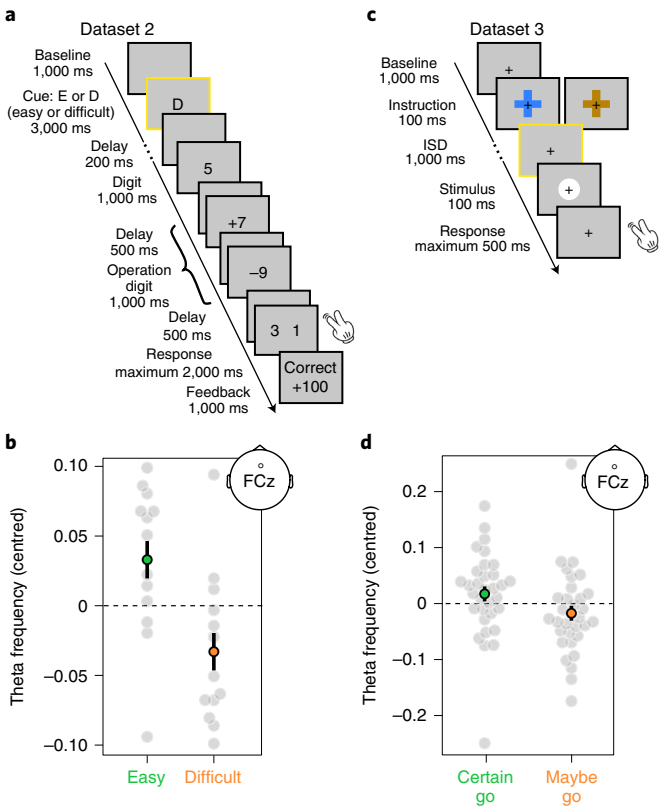


Fig. 5 | Testing model predictions in other datasets. a, Experimental protocol in dataset 2. In this study, participants performed a cued serial mental calculation task in which a cue indicated whether the following calculations would be easy (that is, only +1 additions were used) or difficult (that is, addition or subtraction of different numbers). **b**, Theta peak frequency at FCz for easy and difficult cues. Circles represent correct trials. Peak theta frequency was lower in correct trials in the easy compared with difficult condition (one-sided Wilcoxon sign-rank test: $W = 86$, $P = 0.018$, $r = 0.63$, 95% CI 0.01–0.14). **c**, Experimental protocol in dataset 3. In this study, participants reacted to an action signal (stimulus; white square, circle or triangle). Each shape was randomly assigned towards one action (go, no-go or switch-go). Only the go and no-go actions were analysed, because switch-go data were unavailable. On half of the trials, a certain-go cue (for example, brown cross) indicated that the stimulus was going to be a go shape with 100% certainty. On 50% of the trials, a maybe-go cue (for example, blue cross) indicated that the stimulus had a 25% chance of being a no-go shape. **d**, Theta peak frequency at FCz for certain-go and maybe-go cues. Circles represent correct trials. Peak theta frequency was lower in correct trials following certain-go cues compared with maybe-go cues (one-sided Wilcoxon sign-rank test: $W = 358$, $P = 0.039$, $r = 0.35$, 95% CI −0.01 to 0.08). The yellow framing around the cue and instruction stages in **a** and **c** represents the time window from which we extracted peak theta frequency. Data presented as mean values. Error bars represent s.e.m. computed over $n = 14$ participants in dataset 2 or $n = 33$ participants in dataset 3. Grey dots represent individual participants' data.

views on theta oscillations highlight the importance of peak oscillatory frequency. One exciting avenue of research concerns the interplay between theta oscillations supporting cognitive control and supporting working memory. Indeed, the control over memorized items in a working memory task, for instance after retro-cuing a subset of maintained items or during manipulation of a memorized sequence of items, has been shown to depend on midfrontal theta oscillations[37,38], which are plausibly homologous to the ones observed in the current experiment and in cognitive control more generally[1,39]. Finally, some studies have demonstrated that midfrontal and hippocampal theta oscillations can phase-lock or exhibit coherence with each other in certain contexts[39,40], suggesting that the two theta-generating systems can interact. For now, more studies are needed, for instance using intracranial recordings in humans, to better understand the relationship between midfrontal and hippocampal theta oscillations.

In our model, the MFC unit generates theta oscillations when a rule is instructed. These oscillations orchestrate rule-node competition and generate bursts synchronizing rule-relevant sensory and action nodes. This mechanism based on theta oscillations is coherent with an energizing, or more generally modulatory, role of the dorsal anterior cingulate cortex (dACC), in line with the expected value of control theory[41,42], according to which the dACC specifies the intensity of the control signal. This has also been described as a motivational function of the dACC, in line with the observation that lesion of dACC can lead to deficits in motivated behaviour[43]. However, these accounts do not discuss the specific role or importance of theta oscillations. Another line of work demonstrated that the exertion of cognitive control critically relies on theta oscillations[1,2] to create task-relevant functional networks[3,4]. Most of these studies showed that the amplitude of theta oscillations generated in the dACC increases after conflicts and errors and predicts improvement in task performance[1]. Our study thus extends our knowledge on the energizing role of dACC by showing that, in addition to theta amplitude, another dimension of theta oscillations is crucial for optimal control of task representations, that is, theta frequency. This generalization of the energizing role to a modulatory one allows for an extra degree of freedom in control. Specifically, it posits that two separate aspects of this control signal can be independently manipulated (by the dACC): the intensity of the control signal through theta amplitude and the time window of the control signal's effect on the task representations through theta frequency. Previous studies have shown that adaptive changes in theta amplitude (that is, the intensity of the control signal) are critical for cognitive control, for example, in conflict adaptation[44,45]. On the other hand, adaptively changing the processing time window through theta frequency allows to adjust a trade-off in the orchestration of task representations by theta oscillations. In easy task rules, a faster theta frequency is optimal as these representations are set up quickly and reliably, whereas a slower theta frequency is necessary for difficult rules. This observation opens new avenues for research to understand the functional role of both theta amplitude and frequency in dACC.

Prior models have ascribed some aspects of rule or action representations to the dACC (for example, refs. [46,47]), and it has been shown empirically that dACC represents certain aspects of task sets[41,48–51]. It could thus be argued that our MFC unit should represent some aspects of actions or task sets. Here, we underline that our anatomical labelling (for example, LFC and MFC) was rather broad, in part because the functional architecture is not fully known, especially with respect to the division of labour between lateral and medial prefrontal cortices. Thus, the two theories are not necessarily in contradiction and may simply highlight different functional roles of LFC and MFC. We believe that it will be critically important for future modelling studies to investigate how the modulation of task representations can be implemented through targeted theta oscillations and that future experimental work should aim at disentangling how dACC and LFC modulate and represent task information.

Theta oscillations have also been shown to support attention in a several previous studies (for example, refs. [20,52–56]). Many studies have shown that the amplitude of theta oscillations increases when attention is endogenously oriented (sometimes referred to as sustained attention[52,55]) or when it needs to be reoriented[20,53]. These studies report changes of amplitude in neural oscillations in the theta frequency band or theta band fluctuations in behavioural performance[55–58] (see also ref. [21] for a review). It has been proposed that

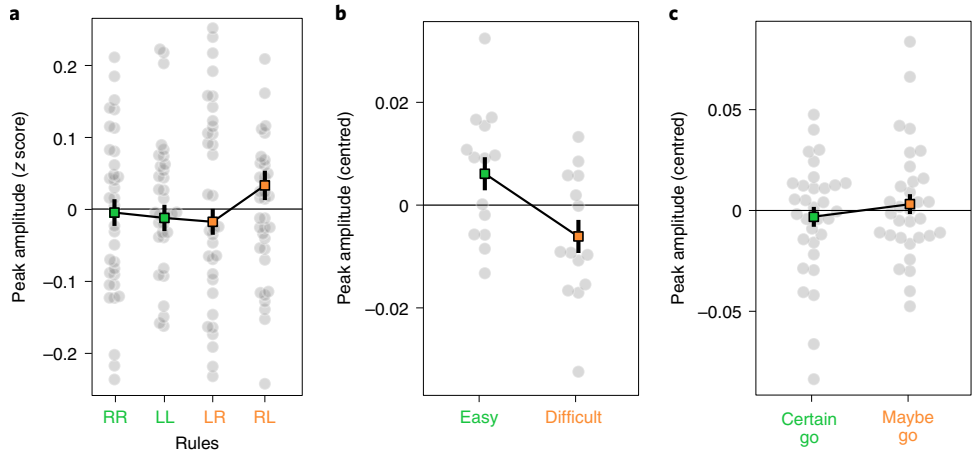


Fig. 6 | Peak theta amplitude per condition in each dataset. a, Peak theta amplitude by condition for correct trials in dataset 1. There was no statistically significant main effect or interaction ($n = 34$ participants, two-way repeated-measure ANOVA: all $F(1, 33)$ values <2.21, all P values >0.146). **b**, Peak theta amplitude by condition for correct trials in dataset 2. There was no statistically significant difference between conditions ($n = 14$ participants, two-sided Wilcoxon sign-rank test: $W = 24$, $P = 0.079$, $r = 0.54$, 95% CI −0.01 to 0.03). **c**, Peak theta amplitude by condition for correct trials in dataset 3. There was no significant difference between conditions ($n = 33$ participants, two-sided Wilcoxon sign-rank test: $W = 257$, $P = 0.896$, $r = -0.02$, 95% CI −0.02 to 0.02). Data presented as mean values. Error bars represent s.e.m. Grey dots represent individual participants' data.

theta oscillations supporting attentional processes reflect rhythmic sampling of visual information at the attended location and across the visual field. These theta oscillations are thought to be supported by a network comprising the lateral intraparietal cortex, the pulvinar nucleus of the thalamus and the frontal eye fields[59,60] (but see also evidence that interaction between local receptive fields in V4 can induce theta oscillations[61]). But to date, none of these studies have reported a shift in theta frequency across conditions. One possibility is that such shifts have been overlooked due to averaging of spectral amplitude across frequencies in the theta frequency band that is commonly performed to test for difference in amplitude of theta band oscillations, or contamination of peak frequency by other factors (for example, the aperiodic component of the spectrum).

However, it seems unlikely that differences in attentional demands underlie our findings. In dataset 1, the difficulty between same-side and different-side instructions (which induced the theta frequency shift) was situated at the stimulus–action mapping level, in contrast with attentional demand manipulations of stimulus discriminability or identity (for example, simple feature versus feature conjunction searches in ref. [57] or the number of stimuli to track in ref. [58]). Indeed, the tilt of grating stimuli in dataset 1 was determined in a separate block before the main experiment and kept constant throughout the main experiment blocks. Furthermore, there was no difference in the validity of rule instructions relative to target location (that is, all instructions were 100% valid) and thus no uncertainty in stimuli location that would differentially affect sampling of visual information by attentional processes. Similarly, no differences specific to attentional orientation or attentional sampling demands distinguished the conditions in datasets 2 and 3. Thus, differences in attentional demands alone cannot explain the shift in theta frequency observed here. Nevertheless, it is possible that different theta oscillation-generating systems coexist and interact to support attention and cognitive control. In fact, in our model, the activity of rule-relevant sensory nodes oscillates at theta frequency due to the bursts sent from the MFC unit. It would therefore be interesting in future studies to investigate how midfrontal theta oscillations supporting cognitive control interact with other generators of theta oscillations shown to support attentional processes[62].

Our study also provides a potential explanation for the discrepancy in reported oscillatory frequencies contributing to top-down cognitive processes[63]. Indeed, several studies have reported the involvement of different low-frequency bands[14,64–67] during top-down control processes (for example, decision-making, working memory and hierarchical task implementation). The intrinsic frequency range of theta oscillations poses limits on the processes that they can orchestrate. Therefore, based on the overwhelming evidence that theta oscillations support task rule implementation and action monitoring[1,68,69], it would seem that such processes must take place within a theta cycle. In our model, we chose to use the canonical 4–7 Hz limits of theta oscillations, thus a task rule that would require a longer build-up time than the slowest theta frequency (4 Hz, period 250 ms) could not be reliably instantiated. However, several recent studies have shown that the implementation of complex task rules (for example, multiple simultaneous novel rules or nested task rules) elicit slow frontal midline oscillations in the delta range (1–3 Hz; see, for example, refs. [65,66]). Although we did not find such a spread into lower or higher frequency bands in our own datasets (Extended Data Fig. 3), these results potentially question the conventional frequency limits of oscillations supporting cognitive control (usually attributed to the theta band). More abstract task rules (see, for example, refs. [65,70]) recruit a larger extent of frontal areas[71] and more rostral frontal areas, in line with accounts of a hierarchical organization of the frontal cortex[70,72]. This larger network of areas might thus require longer periods to build up stable task representations and necessitate slower oscillations (in the low theta or delta bands) to efficiently and reliably implement such abstract task rules.

It would therefore be of interest to further test whether a gradual increase in task complexity or abstractness could elicit a slowing of neural oscillations generated by MFC towards the delta range. For instance, based on the stimulus–action mapping task we developed for our model and dataset 1, it would be interesting for future research to see whether we observe further slowing of midfrontal theta oscillations (that is, into the delta frequency band) if we increase the number of response options from two (that is, CW or CCW) to three or four (that is, different angles of grating rotation relative to vertical). Another possibility to study the effect of gradual increase in task complexity would be to use multistep tasks, such as hierarchical and/or temporally extended tasks[73,74]. Such studies would inform our understanding of the interaction between task complexity and the flexibility of the temporal scale of neural operations.

Our model predicts that a decrease in MFC theta frequency from easy to difficult tasks is beneficial for behavioural performance. Several studies have tested the causal role of theta oscillations in cognitive control using tACS in the theta frequency band[75–77]. In these studies, a fixed theta frequency (for example, 6 Hz) is used across participants and conditions. It would therefore be interesting to test this prediction from our model by varying tACS frequency across the theta frequency range.

The integrator unit in our model aggregates inputs from the action unit, which is itself activated by the sensory unit. The integrator unit commits to a response when the activity of any node in this unit reaches a (collapsing) threshold (Methods). The slope of the information accumulation of the correct node (for a particular trial) thus reflects the strength of the sensory signal or the difficulty of the instructed rule, or more generally, the task difficulty. This relationship with task difficulty is consistent with the relationship observed empirically between the P3b event-related potential component (also referred to as centroparietal positivity) and the difficulty of perceptual decisions[78]. Indeed, the integrator unit of our model constitutes a leaky competing accumulator network[18], which is thought to capture essential dynamics of perceptual decision-making as studied in ref. [78]. One interesting avenue for future research is the observation that our model accumulation in the integrator node associated with the correct response is locked to gamma oscillations that modulated the activity of rate code neurons in each neural triplet. It has been previously shown that evidence accumulation is subject to slow rhythmic fluctuations in the delta frequency band[79]. It would therefore be interesting to test whether additional fast rhythmic dynamics exist in the upslope of the P3b component, which could have been hidden in previous studies. Indeed, if these gamma oscillations were not phase-locked across trials, the event-related potential averaging procedure would average out such fluctuations.

We observed a large interindividual variability in peak theta frequency across conditions (see individual spectra in Supplementary Figs. 3–6). Although interindividual variability of the absolute theta peak (in hertz) can be partly attributed to nonfunctional sources of variance such as skull conductivity and thickness[80,81], variations in individual peak frequency correlate with cognitive performance in the alpha band[82–84] and this variability is related to properties of corticothalamic white-matter projections[85]. This suggests that individual peak frequencies in different frequency bands can be a stable neurophysiological trait[86] and that this variability of peak theta frequency would itself be an interesting topic of investigation for future studies. For instance, investigating the causes and consequences of individual peak theta frequency could have an important impact on the development of personalized neurostimulation interventions using transcranial magnetic stimulation or tACS[77]. Indeed, targeting peak theta frequency could allow to optimally modulate functional connectivity, which has been shown to be dysregulated in Alzheimer's disease[87–89]. Moreover, substantial response variability exists in repetitive transcranial magnetic stimulation treatment using intermittent theta burst stimulation for treatment of major depressive disorder[90]. It would thus be interesting to test whether individualized intermittent theta burst stimulation frequency, estimated in a separate experimental procedure (see, for example, ref. [91]), could, at least partly, reduce this response variability.

Despite the robust and replicable association of theta oscillations and cognitive control, the neurobiological underpinnings of theta generation and modulation remain unclear. Microcircuit models of theta generation in anterior cingulate cortex (ACC) have been proposed[92], and although the relevance of peak frequency fluctuations has been mentioned, no clear mechanism driving such fluctuations has yet been proposed. One candidate mechanism could be a reinforcement-learning system based on ACC–brainstem structures involving the locus coeruleus and noradrenergic neuromodulation of ACC circuits[93]. Indeed, the locus coeruleus heavily innervates MFC and has been shown to modulate cortical oscillations, and its activity increases with task demands[94]. It would thus be interesting to test whether noradrenergic pathways modulate the frequency of ACC-generated theta oscillations in response to task demands. Future studies investigating these candidate neurobiological mechanisms allowing adaptive cognitive control will be crucial to better understand pathogenesis of several psychiatric disorders, for example, attention deficit hyperactivity disorder[93,95].

Neural oscillations may address the fundamental binding problem in cognition by gating information flow in the brain to support cognitive flexibility[33,96]. Our results provide critical insights into the adaptive nature of theta oscillations supporting cognitive control and call for a more systematic evaluation of theta characteristics at computational, behavioural and neurophysiological levels.

Methods

Model. *Overview.* The model implements BC and CTC and consists of five units: two control units (LFC and MFC), two processing units (sensory and action units) and an integrator unit accumulating evidence from the action unit and producing a response. We first briefly describe how BC and CTC are implemented in the model, then proceed to a detailed description of each unit and the nodes composing them.

BC proposes that task representations compete, biased by top-down input. We implemented BC in the LFC unit, which was composed of rule nodes that pointed to specific processing nodes. Each rule node pointed to processing modules composing the rule. This allows a rule node to gate task representations (encoded via an input–output matrix), relevant for that particular rule. For instance, a rule node could implement the rule 'report sensory feature 1 using action set 2' (see this example in Fig. 1b). We used location (L or R) as a sensory feature. We used two action sets, namely L and R hand (see action unit in Fig. 1b). Rule nodes were interconnected to create a competitive accumulator network. Each rule node also received a biasing input throughout a trial from instructions in the form of two letters presented simultaneously and modelled as a top letter instructing which stimulus feature was the target (L or R) and a bottom letter instructing which action set to use (L or R). We refer to these instructions, or rules, in this manner: RL for right–left, where the first letter is the top one, instructing the target stimulus feature (right grating in this case), while the second letter is the bottom one, instructing the action set to use (left hand in this case). Each rule in the task (that is, RR, LL, LR or RL) activated a unique set of instruction nodes (Fig. 1b, LFC unit). Two nodes represented the top letter of an instruction, and two others the bottom letter. This network of instruction nodes created a congruency effect between instruction letters: top and bottom L nodes were connected, thereby activating each other, and similarly for R nodes. In a Stroop-like manner, the connectivity in instruction nodes induced a stronger input to rule nodes for same-side (LL and RR) than for different-side (LR and RL) rules. Furthermore, different-side rules also activated noninstructed instruction nodes more than same-side rules due to the lateral excitation in instruction nodes, thereby making the BC between rule nodes more difficult for different-side rules to win.

The top-down bias signal from control units was implemented through CTC. The MFC unit generated theta oscillations. During a temporal window whose size depended on the specific theta frequency (that is, the slower theta, the longer the temporal window), a competition was initiated between rule nodes. During this competition window, MFC unit sent bursts of activity[11,19,97]. The most active rule node (that is, the one 'winning' the competition) amplified the burst and sent it to sensory and action nodes it points to. All sensory and action nodes oscillated at gamma frequency. These bursts reset the phase of sensory and action nodes selected by the LFC unit, and increased synchrony between them, allowing for efficient communication, that is, gating. Through this selective routing of bursts to sensory and action nodes, the model implements CTC by creating functional networks to implement a rule.

As a result of BC, one rule (typically, the correct one) will win the competition. However, in cases in which the competition is stronger, it will require a longer competition window for the correct rule to win the competition. The latter are difficult rules. In the model, rule difficulty was implemented through conflicting instructions that activated more rule nodes than did easy instructions, making the competition more balanced between the instructed and the other rule nodes. The consequence is that, for difficult rules, the model will achieve better performance at a slower theta oscillation frequency because longer competition will permit rule nodes to win the competition and thus gate the rule-relevant processing nodes. In contrast, for easy rules, performance increases with a higher theta frequency because rule nodes quickly win the competition and the faster theta frequency allows to more frequently gate rule-relevant processing nodes. Hence, an adaptive agent would shift theta frequency depending on task demands.

Oscillatory nodes: a neuronal triplet. In the MFC unit and processing units, each node *i* implements a cortical column simplified as a triplet of neurons, as used in previous models:[19,45,97] a rate code neuron (x_i) and two phase neurons

(one excitatory (E_i) and one inhibitory (I_i) neuron) (Figs. 1b,c and 2b–d). Phase neurons, that is, the E–I pair, generate oscillations with a frequency defined by the E–I pair's coupling parameter (C). This E–I architecture uses the same basic principles as the pyramidal-interneuron network gamma model, which is commonly used to model gamma frequency generation[98–100] but has also been used to simulate neural oscillations in other frequency bands (see, for example, refs. [101,102]). The activity of each phase neuron is defined by a system of stochastic difference equations, following previous work[19,97], for E neurons:

$$E_i\left(t+\Delta t\right)=E_i+\Delta t\left(-CI_i\left(t\right)-\mathrm{Damp}J\left(r>r_{\min}\right)E_i\left(t\right)+B_i\left(t\right)\right), \quad (1)$$

and for I neurons:

$$I_i\left(t+\Delta t\right)=I_i+\Delta t\left(CE_i\left(t\right)-\mathrm{Damp}J\left(r>r_{\min}\right)I_i\left(t\right)\right), \quad (2)$$

in which $E_i\left(t\right)$ and $I_i\left(t\right)$ denote the activity of the excitatory and inhibitory neurons of node i at time t and model data were simulated at 500 Hz, so $\Delta t=0.005$ s. The radius r of oscillation ($r=E^2+I^2$) of an E–I pair, which corresponds to its oscillatory amplitude, is constrained to a radius $r_{\min}=1$ (except for simulations in which we varied MFC theta amplitude; Extended Data Fig. 1c,d). To implement this constraint, we use an indicator function $J(\cdot)$, which returns 1 when its argument is true (that is, when $r>r_{\min}$) and 0 otherwise. The parameter Damp represents the strength of the attraction towards $r_{\min}$ and prevents the activity of the E–I pair from growing too large (that is, it dampens the activity of the E and I neurons). In more neurophysiologically realistic models (see, for example, ref. [103]), such dampening of the E–I pair oscillatory amplitude would be implemented via a projection between the E–I pair and a pool of inhibitory neurons that in turn can inhibit the E and I neurons. For convenience, we here implemented the simpler, approximate implementation via the indicator function J and the Damp parameter, in line with previous models[19,45,97]. The parameter Damp was set to 0.3 for processing nodes. For the MFC node, Damp was set to 0.005 times the theta frequency to scale with the speed of the E–I pair theta oscillations and maintain an equal amplitude across time for all theta frequencies. The term B_i denotes the burst that processing nodes could receive depending on the trial instructions (see MFC unit and LFC unit sections for details). The MFC node did not receive bursts, thus B_{MFC} was set to 0.

The frequency of oscillations generated by the E–I pair was defined by the coupling parameter C, and its relation to frequency in Hertz is given by the following equation:

$$C=f2\pi, \quad (3)$$

in which f denotes the frequency in hertz.

Rate neurons receive, process and transmit information to other nodes. Their activity (x_i) is determined by the input to the node (in_i). For instance, in a sensory node, the input in_i to a rate neuron is either 0 (if its preferred stimulus feature is not presented) or 0.02 (if its preferred stimulus is presented; see Processing units section for more details). This input is then modulated by its excitatory phase neuron (E_i). Thus, rate neuron activity is updated by

$$x_i\left(t+\Delta t\right)=x_i\left(t\right)+\Delta t\left(-x_i\left(t\right)+\mathrm{in}_iF\left(E_i\left(t\right)\right)\right), \quad (4)$$

with $F(\cdot)$ being a logistic function of E_i:

$$F\left(E_i\right)=\frac{1}{1+\mathrm{e}^{-5\left(E_i\left(t\right)-\theta_{\mathrm{E}}\right)}}. \quad (5)$$

Processing units. The processing units are a sensory unit and an action unit. Each unit is composed of nodes representing cortical columns (Oscillatory nodes: a neuronal triplet section).

In all nodes of sensory and action units, the coupling parameter C was set to generate gamma oscillations. The gamma oscillations were set to 30 Hz by using a coupling parameter of $C=188.5$, which in the computational implementation of the model was set to 0.377 to account for the sampling rate of 500 Hz ($C/500=0.377$). We used low gamma-band oscillations around 30 Hz as this subband of gamma has been shown to be critically important for visual processes[104,105] and to be modulated by theta-band oscillations following task cues[11]. To test model stability, we also ran simulations using a higher gamma frequency of 50 Hz, finding similar results.

To further show model stability and induce noise in processing nodes' oscillatory phase, we modified the neural triplet dynamics used in prior implementations[19] in which noise was introduced by independently varying the oscillatory frequency of each neural triplet across trials, while fixing it across time for each single trial. For that purpose, we added random slow fluctuations in the coupling parameter of nodes oscillating at gamma frequency, thereby mimicking noise in ongoing gamma oscillations as observed in empirical studies[106]. We generated random numbers from a normal distribution with parameters $\mu=1$ and $\sigma=1$, for each trial and each processing unit (that is sensory and action). A low-pass filter was then applied to these coupling fluctuations time courses, that is, Gaussian convolution with $\sigma=1$ (in seconds). Finally, the coupling parameter (that is, $C=0.377$ for 30 Hz oscillations) was multiplied by the value of these low-frequency coupling fluctuations. The result of this manipulation was slow random fluctuations of gamma frequencies in phase neurons of processing units. For example, for one trial, sensory nodes were oscillating at 32 Hz at a certain time t, then gradually shifting to 27 Hz, then to 35 Hz, etc. This slow fluctuation was generated independently for sensory and action units.

Rate neurons of the action unit receive input from rate neurons of sensory nodes to implement the two-alternative orientation-discrimination task on gratings. The main task was to report whether the target grating was tilted CW or CCW from the vertical axis. To report the tilt, the rule was to use the index and middle fingers of either the left or right hand, as indicated by the instructions. The left middle finger and right index finger should be used to report a grating tilted CW, and the left index finger and right middle finger to report a grating tilted CCW. Therefore, the connectivity between sensory and action nodes' rate neurons implemented this rule.

Integrator unit. The integrator unit accumulates information for each response, and triggers the model response once one of the integrator nodes reaches a threshold. There is thus one integrator node for each action node. The integrator nodes constitute a competitive accumulator network (as implemented in prior work, for example, ref. [18]) with the following update:

$$\mathbf{y}_{\mathrm{Integ}}\left(t+\Delta t\right)=\mathbf{y}_{\mathrm{Integ}}\left(t\right)+\Delta t\left(\mathbf{W}_{\mathrm{Integ}}\mathbf{x}_{\mathrm{A}}+\mathbf{W}_{\mathrm{lat,Integ}}\mathbf{y}_{\mathrm{Integ}}\left(t\right)\right)+\sigma_{\mathrm{Integ}}\mathbf{N}\left(t\right), \quad (6)$$

in which $\sigma_{\mathrm{Integ}}=0.05$ is a vector collecting the activity of all integrator nodes at time t, $\mathbf{W}_{\mathrm{Integ}}$ denotes the weight matrix between action nodes and integrator nodes and $\mathbf{x}_{\mathrm{A}}$ denotes input from action nodes to integrator nodes. $\mathbf{W}_{\mathrm{lat,Integ}}$ denotes the update matrix of integrator nodes in which off-diagonal cells are set to −0.10 to implement lateral inhibition while diagonal cells, representing the update rate of the competitive accumulator network, are set to 1. Finally, noise was added for each of the four variable integrator nodes with $\sigma_{\mathrm{Integ}}=0.05$ multiplying a vector $\mathbf{N}\left(t\right)$ of four random values drawn from a standard-normal Gaussian distribution.

As stated above, the integrator unit produces a response when a threshold is reached by one of the integrator nodes. To model a speeded task constraint, we modified the classic competitive accumulator network[18] to implement a collapsing threshold, equivalent to a collapsing bound in the DDM, which has been shown to adequately model the dynamics of response threshold in speeded tasks[107]. The threshold θ_{y} therefore decreased exponentially from stimulus presentation to response deadline according to the equation

$$\theta_{\mathrm{y}}\left(t\right)=4-\left(1-\mathrm{e}^{\left(-\frac{t}{0.35}\right)^2}\right)\frac{a}{2}, \quad (7)$$

in which $\theta_{\mathrm{y}}(t)$ denotes the threshold of the integrator unit at time t and a denotes the initial starting point of θ_{y}. In all simulations, a was set to 4. Once one of the four Integrator nodes reached the threshold, we recorded the accuracy, depending on instruction, stimuli and the integrator node which reached the threshold, and the time elapsed from stimuli onset, which provided reaction time for this response (Fig. 2e).

MFC unit. The MFC unit generates theta oscillations that (1) generate bursts that phase-reset the processing units, as in prior work[19,45,97], and (2) additionally initiate a competition window in LFC nodes (Fig. 2a,b).

The MFC unit is composed of a single node in which the E–I pair generates theta oscillations, whose frequency depends on the coupling parameter between the E–I pair. The rate neuron of the MFC node follows a Bernoulli process (Be) with a probability defined by the activity of the node's E neuron as

$$\mathrm{MFC}_x\left(t\right)=\mathrm{Be}\left(\frac{1}{1+\mathrm{e}^{-5\left(E_{\mathrm{MFC}}\left(t\right)-\theta_{\mathrm{burst}}\right)}}\right), \quad (8)$$

In which Be denotes the Bernoulli process, E_{MFC} denotes the activity of the MFC E neuron and θ_{burst} (set to −1) denotes the offset of the relation between E_{MFC} and p (the probability to trigger a burst). Be(p) is 1 with probability p. It will thus typically be 1 when the $E_{\mathrm{MFC}}(t)$ oscillation is near its peak. When $\mathrm{MFC}_x=1$, a fixed amplitude burst of 0.5 is emitted to the LFC unit. The purpose of this burst is to synchronize processing nodes selected by the LFC, by phase-reset of their E neuron (see Processing units section for the burst's effect and LFC unit section for the selection of the processing nodes receiving the burst).

In addition to the burst-emitting function of the MFC proposed in earlier work[3,19], the MFC in the present model opens a competition window between rule nodes in the LFC at each cycle of its theta oscillations. At each cycle of theta oscillations in E_{MFC} activity, a competition window is opened in which LFC rules compete. This competition starts when $E_{\mathrm{MFC}}>\theta_{\mathrm{comp}}$, with $\theta_{\mathrm{comp}}=0.1$. The competition window lasts a fixed temporal interval across cycles defined by (as just defined) θ_{comp} and the crucial C_{MFC} parameter, which determines the theta frequency. To simulate different theta frequencies in the MFC, we varied the MFC coupling parameter (C_{MFC}) from 0.050 (for 4 Hz theta), to 0.087 (for 7 Hz theta). See equation (3) in Oscillatory nodes: a neuronal triplet section.

LFC unit. To implement biased competition in rule implementation, we extended previous models simulating task rules. We considered rule nodes as pointers to processing nodes constituting components of the rule (see, for example, refs. [41,108]). Such pointers permit to bias processing units according to task rules and to create bindings between task-relevant components (see, for example, ref. [109]). In recent computational accounts incorporating oscillations and synchrony, the LFC has been hypothesized to contain such pointers which route MFC bursts to processing nodes[3,19,45,97]. However, in these latter models, no competition occurs between rule nodes. In the current model, LFC is composed of rule nodes, where each such node consisted of one rate code neuron only. Together, they form a competitive accumulator network[18], thereby implementing competition between rules. Each rule node receives a constant input throughout a trial from instruction nodes, which themselves are activated by the two instruction letters. Two instruction nodes represent the top letter of an instruction, and two other instruction nodes the bottom letter. This network of instruction nodes implements a congruency effect between instruction letters: top and bottom L nodes were connected with a positive weight, thereby activating each other, and similarly for R nodes (see instruction nodes in Fig. 1b). Instructions are represented as a vector of binary values (zeros and ones) in which the first two indices represented a top L and R, respectively, and the two last indices represented the presence of a bottom L and R, respectively. For instance, the rule RL was represented as $\mathbf{instructions} = [0, 1, 1, 0]$. This was the input to the instruction nodes, which then projected to rule nodes through the equation

$$\mathbf{in}_{\text{rule}} = \mathbf{W}_{\text{instruction}}\mathbf{instructions}, \quad (9)$$

in which $\mathbf{in}_{\text{rule}}$ denotes the input to rule nodes (that is, from instruction nodes). The matrix $\mathbf{W}_{\text{instruction}}$ represents the connectivity between instruction nodes implementing the lateral excitation, that is, instruction letter congruency effect. The diagonal of $\mathbf{W}_{\text{instruction}}$ was set to 1, and the cells representing the positive weight implementing the lateral excitation were set to 0.5.

The activity of rule nodes is updated through the equation

$$\mathbf{y}_{\text{rule}}(t + \Delta t) = \mathbf{y}_{\text{rule}}(t) + \Delta t\,(\mathbf{W}_{\text{in}}\mathbf{in}_{\text{rule}} + \mathbf{W}_{\text{lat,rule}}\mathbf{y}_{\text{rule}}(t)) + \sigma_{\text{rule}}\mathbf{N}(t), \quad (10)$$

in which $\mathbf{y}_{\text{rule}}(t)$ denotes the activity of all rule nodes at time t, $\mathbf{in}_{\text{rule}}$ denotes the input to rule nodes (that is, instructions) and $\mathbf{W}_{\text{in}}$ denotes the weight matrix between instruction nodes and rule nodes in which weights between an instruction node of a particular letter and rules containing this letter were set to 0.5. For example, instruction nodes 'top R' and 'bottom L' projected to the rule node RL with weight 0.5 (see connectivity between instruction and rule nodes in Fig. 1b). $\mathbf{W}_{\text{lat,rule}}$ denotes the update matrix of rule nodes in which off-diagonal cells are set to −0.1 to implement lateral inhibition. The diagonal cells, representing the update rate of the competitive accumulator network, are set at 0.13. Finally, noise was added for each of the four rule nodes with $\sigma_{\text{rule}} = 0.075$, multiplying a vector $\mathbf{N}(t)$ of four random values drawn from a standard-normal Gaussian distribution.

This architecture from instruction nodes to rule nodes allowed to manipulate task difficulty. For instance, the same-side rule LL, modelled as $\mathbf{instructions} = [1, 0, 1, 0]$, provided strong input to the LL rule node, and a small input to the LR and RL rule nodes as they each share the bottom and top letter, respectively, with the instruction LL. Thus, for $\mathbf{instructions} = [1, 0, 1, 0]$, $\mathbf{in}_{\text{rule}} = [0, 1.4, 0.7, 0.7]$, in which the $\mathbf{in}_{\text{rule}}$ indices represent (in this order) RR, LL, LR and RL. On the other hand, a different-side rule such as RL, modelled as $\mathbf{instructions} = [0, 1, 1, 0]$, provided a relatively strong input to the RL rule node and a small input to LL, RR and LR nodes. Thus, for $\mathbf{instructions} = [0, 1, 1, 0]$, $\mathbf{in}_{\text{rule}} = [0.7, 0.7, 0.4, 1]$, creating a stronger competition between the instructed rule (RL) and the other rules (RR, LL and LR) (Fig. 2a).

Finally, the most activated rule node at each time t amplified and routed the burst emitted at time t by the MFC ($\text{MFC}_x(t)$) to the processing nodes it points to

$$\mathbf{B}(t) = \mathbf{LFC}_{\text{pointers}}\,[\mathbf{y}_{\text{rule}}(t) \circ J(\mathbf{y}_{\text{rule}}(t) = \max(\mathbf{y}_{\text{rule}}(t)))]\,\text{MFC}_x(t), \quad (11)$$

in which $\mathbf{B}(t)$ is a vector of burst values arriving at each processing node's E neuron (to reset its phase). $\mathbf{y}_{\text{rule}}(t)$ is the activity of rule nodes at time t and the circle represents point-wise product. $J(\cdot)$ is an indicator function that returns an array of 0 and 1, with 1 only for the most activated rule node at time t. $\mathbf{LFC}_{\text{pointers}}$ is a matrix containing the processing nodes each rule node is pointing to. $\text{MFC}_x(t)$ is the activity of the MFC rate neuron at time t. This could be 0 or 0.5 (activity values were fixed), depending on whether the MFC is emitting a burst or not at that particular time point. Critically, equation (11) shows that only processing nodes corresponding to the most activated rule node received the burst, while all other processing nodes did not. For instance, if the instructed rule is RR and the most activated rule node at time t is RR, the sensory module 'Right grating' and the action module 'right hand' received the burst, thereby synchronizing their gamma oscillations.

As a result of the congruency in instruction letters and BC between rule nodes, the instructed rule will win the competition more quickly for same-letter rules, that is, easy rules, than for different-letter rules, that is, difficult rules (Extended Data Fig. 1a,b). Therefore, same-side rules will succeed in synchronizing rule-relevant processing nodes more quickly. One consequence is that, for difficult rules, the model will achieve better performance at a slower theta oscillation frequency when competition lasts longer. In contrast, for easy rules, model performance increases with a slightly higher theta frequency. Hence, an optimal agent would shift theta frequency depending on task demands.

Simulations. We ran simulations of the model on the task depicted in Fig. 1a. Instructions are shown for 200 ms (two letters), then a variable ISD between 1,700 and 2,200 ms in 11 steps of 50 ms allows to prepare the instructed mapping, and subsequently two gratings are shown for 50 ms. There were four possible instructions: RR, LL, LR and RL.

The presentation of each possible stimulus was modelled as constant input set to a value of 0.02 to the corresponding sensory node. There were four possible stimulus configurations because each of the two gratings could be tilted either CW or CCW. For each combination of task parameters, we ran 100 repetitions, which amounts to (11 ISD + 4 instructions + 4 stimuli configurations) × 100 = 17,600 trials. We then grouped repetitions into 34 groups of ~500 trials each, each representing one participant.

Effect of amplitude on competition window. To verify that high and low theta frequencies are optimal for easy and difficult tasks, respectively, we independently varied MFC theta amplitude and frequency, and computed the competition window lengths for each combination (Extended Data Fig. 1c). Higher theta amplitudes increased the competition window length but quickly reached a ceiling (around an amplitude of 3). Theta frequency on the other hand produced larger increases in competition window, indicating that effects of theta frequency on model performance cannot be explained by theta amplitude alone. Furthermore, we replicated our main simulation at different MFC theta amplitudes and obtained similar results (that is, the difference in optimal frequency for easy and difficult rules; Extended Data Fig. 1d).

Stimulus–action mapping experiment (dataset 1). *Participants.* Thirty-nine human participants were recruited for this experiment (mean ± s.d. age 23.7 ± 4.5 years, range 18–41 years old; 27 female). All participants had normal or corrected-to-normal vision and no history of neurological problems. All participants provided written informed consent and received monetary compensation for their participation. Five participants were excluded from the analysis: two completed fewer than 5 blocks, one had fewer than 200 trials after trial rejection based on eye-tracking data, one had poor overall behavioural performance (that is, less than 50% overall accuracy) and one participant was left-handed. The experiment was approved by the local ethics committee (Faculty of Psychology and Educational Sciences, Ghent University). Sample size was not computed a priori: we aimed for more than 30 participants. First, we recruited 35 participants to reach a total number of more than 30 participants after drop-out, considering a ~10% drop-out rate due to noise-corrupted data or other issues related to participants' task performance. The sample size after exclusions dropped to 30 participants, thus we tested 4 more participants, who were all included, bringing the sample size to 34 participants. Assuming a medium effect size and aiming for a power of 0.8 in a within-subject repeated-measure ANOVA analysis, the study would require a sample of 32 participants. Data collection and analysis were not performed blind to the conditions of the experiments.

Apparatus and stimuli. Participants sat in a dimly lit room, 60 cm from a 24 inch liquid-crystal display monitor (refresh rate 60 Hz, resolution 1,280 × 1,080 pixels). A chinrest was used to stabilize head position and distance from the screen. The experiment was implemented using Python 2.7 and the PsychoPy toolbox[110].

Experimental design. Participants were instructed to perform a two-alternative forced-choice orientation-discrimination task on two sinusoidal gratings presented simultaneously on each side of a central fixation cross (Fig. 1a). Each grating was randomly tilted either CW or CCW relative to the vertical axis. The stimuli were sinusoidal gratings windowed by a raised cosine (size of 5° visual angle, 10% contrast, three cycles per degree, at 5° eccentricity, on a grey background). The tilt angle was calculated for each participant using a staircase procedure (see below) to avoid ceiling accuracy. Participants were instructed at the beginning of every trial to perform the two-alternative forced-choice task on the right or left grating and to respond using their right or left hand (index and middle finger, respectively, for CW and CCW tilt).

Instructions letters were presented for 200 ms with a size of 0.75° visual angle, and positioned above and below the central fixation cross (vertical eccentricity of 1° visual angle). The letter above the fixation cross instructed which grating was the target, that is, on which grating the discrimination should be performed, and the letter below the fixation cross instructed which hand to use to respond. After instructions, a preparation interval followed to allow participants to process instructions and prepare the stimulus–action mapping to perform the task. We used a dense behavioural sampling paradigm with multiple, densely distributed ISD:[20] the duration of the ISD, between instructions and stimuli, was randomly chosen on each trial from 11 possible durations from 1,700 to 2,200 ms in 11 steps of 50 ms. The variation in ISD was introduced to measure oscillations in behavioural performance and test predictions of the model (Fig. 3c).

A trial time course consisted of a 1,000 ms baseline period, followed by instruction presentation for 200 ms, then the ISD, and finally the stimuli presentation for 50 ms. After stimuli onset, the fixation cross turned blue, indicating the beginning of the 700 ms response window. If a correct response was given, the fixation cross turned green. If an incorrect response was given, the fixation cross turned red. If no response was given during the response window, a message indicated that the participant was too slow and the experiment was paused, prompting the participant to take a break if needed, and press 'Space' to resume the experiment. Every trial that was missed, that is, not responded to, was added to the trial queue, and presented again at the end of the block. Participants performed one training block to familiarize themselves with the experimental design, one staircase block to compute the participant's grating tilt angle and between five and eight blocks of the task depending on the number of missed trials (that is, participants who missed more response deadlines had longer blocks because trials were queued at the end of the block). The practice block consisted of 80 trials, the stimulus was shown for 100 ms and the response window lasted 1,000 ms to make the practice task easier.

Following the practice block, participants completed a block implementing a staircase procedure on the tilts of the gratings. The staircase was done across all instructions and all ISDs to find a tilt level that would avoid ceiling performance and thus allow for variability across ISDs. We used a one-up two-down staircase procedure consisting of 80 trials. The event timings and stimulus properties were the same as in the main task. Only the tilt of the gratings varied throughout the trials. Initially, a wide tilt (7°) was set. The procedure started with a step size of 3°, which was divided by 2 every other reversal starting at the second reversal. The reversal corresponded to switches in participants' response accuracy, that is, from a sequence of correct responses to an incorrect response or the other way around. When a participant switched from a correct response to an incorrect response, the difficulty of the task was decreased by increasing the tilt of the gratings. Conversely, when a participant responded correctly after a sequence of errors, the difficulty of the task was increased, that is, the tilt of the gratings was decreased. The minimum tilt step size was set at 0.1°, the maximum final tilt of the gratings was 30° and the minimum was 0.5°. The final tilt was the average of the last ten tilts.

After the staircase block, participants completed between five and eight blocks of the main task depending on the number of trials missed. That is, participants who missed the response deadline more often had longer blocks (because of queued trials) and therefore completed fewer blocks. In total, the experiment lasted ~3 h from explanation of the task to removing the EEG cap.

Eye-tracking acquisition and processing. We recorded eye movements using a SensoMotoric Instruments RED250 mobile system eye-tracker with a sampling rate of 250 Hz. The eye-tracker camera using infrared optics was attached to the bottom of the computer screen. We used the PyGaze Python toolbox[111] to control the eye-tracker through the experiment's script. Each block of the experiment started with a calibration procedure in which participants had to follow a moving red dot with their eyes to nine locations on a grey background, the success of which was validated before continuing. Gaze position was epoched from instructions onset to stimulus presentation. To epoch gaze position data and align them with EEG data, we aligned the trial onset (instructions presentation) using the trial onset trigger in eye-tracking data and the trial onset trigger in EEG data. We then calculated the distance from the fixation cross in degrees of visual angle at each time point in the epoch. Any trial in which the gaze was outside a 1.5° radius centred on the fixation cross at any moment in the ISD was rejected in the behavioural and EEG data.

Behavioural data analysis. As described above, trials in which gaze position distance from the fixation cross exceeded 1.5° of visual angle were discarded. Trials were grouped by instruction and by ISD. Model simulations showed a theta frequency–rule difficulty interaction in accuracy but not in reaction times (Fig. 3a and Extended Data Fig. 2a). We therefore used accuracy as our dependent variable.

To compute spectra of behavioural accuracy oscillations across ISDs, we first average-padded accuracy values (Supplementary Fig. 2a). Average-padding was performed for each participant and for each instruction independently to increase frequency resolution to 1 Hz (ref. [20]). To pad the data, values corresponding to average accuracy across ISDs (by instructions) were added on either side of the empirical data points. Specifically, the eleven time points, spanning 500 ms, were padded to get a 1,000 ms segment, thus adding five data points before the first data point and five after the last one.

We then computed a fast Fourier transform to obtain frequency spectra of each accuracy-by-ISD time course for each participant and each instruction. Fast Fourier transform allows to decompose the behavioural data from the time domain into frequency components to estimate an amplitude spectrum, that is, the amplitude of oscillations at each frequency present in the original data. We then extracted peak theta frequency by selecting the frequency with the largest amplitude. Finally, we *z* scored the peak frequency value across rules, separately for each participant, to discard any difference in offset or range of theta peak frequencies across participants. This procedure was carried out to specifically test the model prediction that theta peak frequency decreases with task difficulty, thus interindividual differences in theta peak frequency for each instruction were not of interest in this specific analysis. The raw peak frequencies for behavioural oscillations are also available in Supplementary Fig. 7a (left panel).

EEG acquisition and preprocessing. EEG was recorded using a Brain Products actiChamp system with 64 active scalp electrodes positioned according to the standard international 10–20 system at a sampling rate of 512 Hz. Four electrooculographic channels were used to record eye movements and blinks: two were placed on the outer canthi of the eyes, and two were placed above and below the right eye. All preprocessing steps were carried out by using the Python MNE toolbox v.0.21[112]. Raw EEG data were downsampled offline to 200 Hz, re-referenced to the average reference and low-pass filtered at 48 Hz using a finite impulse response filter with a Hamming window. The analysis of the prestimulus interval was performed on epochs from −1,000 to 0 ms relative to stimulus onset, yielding epochs of 1,000 ms. A linear detrend was performed on each epoch individually. After trial rejection based on eye-tracking data (Eye-tracking acquisition and processing section), raw EEG and electrooculographic time courses were visually inspected on a trial-by-trial basis to reject visible artefacts, eye movements or blinks. The mean ± s.d. percentage of rejected trials across participants was 26 ± 14%.

EEG spectral analysis. To estimate peak frequency of theta oscillations, we first computed power spectral density over the 1,000 ms window using Welch's method provided in the Scipy toolbox v.1.3.1[113]. The Welch power density estimation was performed using a Hann window and zero-padding to obtain 400 time points of data to smooth the spectra to improve estimation of peak frequency in the following analysis step. We then used a recent method that allows to parametrize neural spectra by fitting the $1/f$ pattern in electrophysiological recordings spectra (also called the aperiodic component), and subsequently identifies spectral peaks by fitting Gaussians on the flattened spectrum (that is, after removing the aperiodic component). This method thereby provides a sensitive identification and estimation of oscillatory processes in neural activity (FOOOF toolbox, version 1.0.0[16]).

Indeed, this method permits to deconfound several factors that can mask shifts in peak theta frequency in grand average spectra (Supplementary Fig. 3a). First and foremost, interindividual differences in the $1/f$ structure (also called aperiodic component) of the spectrum can mask shifts of peak theta frequency across conditions. More specifically, the offset and slope (also called exponent) of the aperiodic component have been shown to vary across participants[16,114–116] (Supplementary Fig. 3b). This variability can therefore affect the apparent peak frequency in the grand average spectrum. Second, interindividual variability in the height of theta band peaks makes it more difficult to compare the grand average (as can be seen in the grand average spectra in Supplementary Fig. 3a). And third, relatively large peaks in the alpha band (which can be as much as six times larger in power than the theta frequency peaks in some participants) vary in peak frequency and width across participants (see Supplementary Fig. 3b (middle panel), individual participants' spectra in Extended Data Fig. 3 and Supplementary Figs. 5 and 6). These large peaks in the alpha band can alter the shape of the grand average spectra and mask changes in peak theta frequency. Together, these confounding factors require the estimation of the aperiodic component of the spectrum and the independent estimation of oscillatory peaks over the aperiodic component, as is performed in the FOOOF toolbox[16] (but see below for control analyses in which we show that our main results are visible in raw power spectra, and robust and statistically significant when estimating theta peak frequency based on the raw power spectra, that is without the FOOOF toolbox).

This algorithm yields several measures, including the peak frequency and amplitude of oscillations detected over the $1/f$ pattern in the spectra (that is, by reporting the mean and height of the Gaussian fitted to each identified spectral peak in the flattened spectrum; Supplementary Fig. 2b). Using this algorithm, we computed, separately for every participant, trial and electrode, whether a peak was detected in the theta frequency range (that is, higher than 3 Hz and lower than 8 Hz), and we saved the estimated peak (in hertz) and the amplitude of the peak (in $\mu V^2 Hz^{-1}$). Settings for the FOOOF algorithm were as follows: To obtain peak frequency and amplitude in the theta frequency range, the power spectra were parametrized across the frequency range from 2 to 20 Hz. The peak width limits were set between 0.5 and 2, to find peaks that were frequency specific. The maximum number of peaks was set at four, under the assumption that, in the 2–20 Hz frequency range, there could be four meaningful peaks, that is, one in each band (delta, theta, alpha and beta). No minimum peak height was set, the peak threshold was set at 2 (default) and the aperiodic mode was fixed (default).

To test model predictions in theta peak frequency, we separated trials according to the instruction and accuracy for each participant and each electrode. As a sanity check, for replication of previous findings on theta amplitude and as an independent electrode selection procedure, we investigated the scalp distribution of theta oscillation power for correct versus incorrect trials for dataset 1 (FCz was selected based on previous findings[1] for datasets 2 and 3 since there were not enough incorrect trials; Results). To perform this analysis, we extracted the power of peaks found in the theta band using the FOOOF toolbox, for each participant, for each electrode, for each instruction, and for each trial separately. We then *z*-scored these values across electrodes. This allowed to highlight the specific theta

power topography elicited by proactive cognitive control, that is, preparing to implement an instructed stimulus–action mapping. We performed a cluster-based permutation test[117] with 10,000 permutations on scalp topographies to test whether a cluster of electrodes showed relatively higher theta power in correct versus incorrect trials (across all instructions). This analysis revealed a statistically significant cluster of electrodes in frontocentral sites (permutation cluster test $P < 0.001$; Fig. 4c). We then computed the average peak theta frequency (in hertz), extracted using the FOOOF toolbox, in the selected cluster of electrodes (Fig. 4c). In dataset 1, we *z*-scored the peak frequency value across rules, separately for each participant, to discard any difference in offset or range of the EEG theta peak frequencies across participants. In datasets 2 and 3, we centred the peak frequency value across conditions, separately for each participant, for the same reasons and because there were only two conditions. This procedure was carried out to specifically test the model prediction that theta peak frequency decreases with task difficulty, thus interindividual differences in theta peak frequency for each instruction were not of interest in this specific analysis. The raw peak frequencies for EEG oscillations in all three datasets are also available in Supplementary Fig. 7a–c.

Because of the challenging nature of characterizing power spectra (see, for example, ref. [118]), we also carried out additional control analyses directly on the raw power spectra to test whether our main result was observable without using the FOOOF toolbox (that is, without aperiodic component estimation and Gaussian fitting). First, we sought to better illustrate the shift in theta peak frequency that can be confounded by interindividual differences and other factors (Supplementary Fig. 3a). To do so, we replotted the grand average spectra (from Supplementary Fig. 3a) after aligning each participant spectrum to its own peak in the most difficult rule (that is, RL) and scaling each grand average spectrum between 0 and 1. This analysis, anchored in the raw spectra (without using the FOOOF toolbox) showed that a clear shift towards higher frequencies was visible for easier rules, confirming that our results are robust and observable at the group level.

Second, we showed that our main finding that frontal peak theta frequency decreases with rule difficulty is also present (both substantially and statistically) when estimating theta peak frequency directly in the raw amplitude spectra as the theta-band frequency exhibiting the highest power (in $\mu V^2 Hz^{-1}$; Supplementary Fig. 8). More specifically, for each participant, condition and electrode in the identified frontocentral cluster, we extracted the theta-band frequency with the highest power, then averaged these peak frequencies across electrodes, resulting in a theta peak frequency per participant per condition. Moreover, to evaluate the presence of our predicted effect, that is, a decrease in peak theta frequency from same-side to different-side rules, we computed the proportion of participants exhibiting a positive difference in peak theta frequency between same-side and different-side rules. We carried out this analysis in both peak frequencies obtained using the FOOOF toolbox and peak frequencies obtained directly from the raw spectra. This analysis revealed that a majority of participants exhibited the predicted effect (see Supplementary Fig. 8 and discussion thereof), showing once again that our results are replicated without the FOOOF toolbox, thereby alleviating any concern about the challenges in the estimation of the aperiodic component performed by the FOOOF toolbox[118]. However, we still believe that condition- and participant-wise estimation of the aperiodic component of the spectra, and frequency and power of peaks (as performed by the FOOOF toolbox) is necessary to avoid confounding factors that could mask the effect or result in spurious differences in peak frequency or power.

Finally, for the control analyses on theta amplitude (Fig. 6) and nearby frequency bands (Extended Data Fig. 4), we followed the same procedure as for the main analyses on theta frequency band (that is, using the FOOOF toolbox). For the alpha frequency band, we followed the same procedure as for the main results on the theta band but considered peaks in the 8–12 Hz range. For the delta frequency band, we used frequency limit parameters of [0, 20] Hz in the parametrization of the power spectra, and considered peaks in the range of 1–3 Hz. For both delta and alpha frequency bands, we then followed the same procedure as for the main results on the theta frequency band.

Statistical analyses. To compute the optimal theta frequency per rule difficulty (same-side (RR, LL) versus different-side (LR, RL) rules) in the model, we calculated the MFC theta frequency yielding the highest accuracy for each group of simulations (Simulations section). We then compared the two samples of optimal theta frequencies per rule difficulty using a two-sided Wilcoxon signed-rank test (Fig. 3b). For this, and all other Wilcoxon signed-rank tests, we computed nonparametric confidence intervals using bootstrapping of the difference in medians between conditions. The reported effect size for Wilcoxon signed-rank tests is the matched-pairs rank-biserial correlation (r)[119].

Reaction times and DDM parameters estimated on model data (Extended Data Fig. 2a) were analysed using a 2×7 repeated-measure ANOVA with factors of rule difficulty (two levels: same-side and different-side) and theta frequency (seven levels: from 4 to 7 Hz in steps of 0.5 Hz), using the StatsModels v0.10.1 (https://www.statsmodels.org/v0.10.1/) and Pingouin v0.5.0[120] (https://pingouin-stats.org/) Python packages. Data distribution was assumed to be normal, but this was not formally tested.

Participants' behavioural and EEG data from the stimulus–action mapping experiment (dataset 1) were entered into two-way repeated-measure ANOVAs with factors of target location (two levels: L and R) and hand (two levels: L and R) using the StatsModels and Pingouin Python packages. For behavioural data, this consisted of accuracies, reaction times and DDM parameters per rule (Fig. 4a and Extended Data Fig. 2a) and peak theta frequency (that is, of accuracy-by-ISD) per rule (Fig. 4b). For EEG data, this consisted of the average peak theta frequency from the selected electrode cluster (Methods) of correct trials per rule (Fig. 4d, circles), and the difference between average peak theta frequency of correct and incorrect trials.

For datasets 2 and 3, we report the statistical results from their behavioural data analyses. For the analysis of their EEG data, we followed the same procedure as for dataset 1, and because only two difficulty levels were available in these datasets, we performed a one-sided Wilcoxon signed-rank test to test the hypothesis that theta frequency decreased from an easy to a difficult condition. For all control analyses, that is, effect of task difficulty on peak theta amplitude (Fig. 6), and on peak frequency and amplitude in the delta and alpha frequency bands (Extended Data Fig. 4), we performed two-sided Wilcoxon signed-rank tests.

To investigate interindividual differences in the sensitivity of EEG peak theta frequency to rule difficulty in dataset 1, we performed a linear regression of each participant's raw EEG peak frequency (in hertz) in correct trials ordered by each rule's overall accuracy across participants (that is, rule was treated as a linear predictor with 79.16% for RR, 76.70% for LL, 71.87% for LR and 71.44% for RL). In a second step, individual-participant slopes were correlated with overall accuracy, collapsed across rules (Fig. 4e). We used a robust Spearman correlation (that is, skipped correlation[121]) implemented in the Pingouin Python package, which identifies outliers based on the minimum covariance determinant. The effect was also significant using a simple Spearman correlation. For illustration purposes, we computed a linear regression, excluding the identified outlier, to plot it as a regression line in Fig. 4e.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Raw behavioural, eye-tracking and EEG data can be found on the Open Science Framework repository at https://osf.io/nwh87/?view_only=b11ee1f860804da582c816fe8acdecad.

Code availability

Code of the model, the behavioural experiment and analysis scripts to reproduce all results and figures from the study can be found on Github at https://github.com/mehdisenoussi/theta_shift_cog_control.

Received: 5 April 2021; Accepted: 10 March 2022;
Published online: 21 April 2022

References

1. Cavanagh, J. F. & Frank, M. J. Frontal theta as a mechanism for cognitive control. *Trends Cogn. Sci.* **18**, 414–421 (2014).
2. Sauseng, P., Tschentscher, N. & Biel, A. L. Be prepared: tune to FM-theta for cognitive control. *Trends Neurosci.* **42**, 307–309 (2019).
3. Voloh, B. & Womelsdorf, T. A role of phase-resetting in coordinating large scale neural networks during attention and goal-directed behavior. *Front. Syst. Neurosci.* **10**, 18 (2016).
4. Canolty, R. T. et al. High gamma power is phase-locked to theta oscillations in human neocortex. *Science* **313**, 1626–1628 (2006).
5. Varela, F., Lachaux, J.-P., Rodriguez, E. & Martinerie, J. The brainweb: phase synchronization and large-scale integration. *Nat. Rev. Neurosci.* **2**, 229–239 (2001).
6. Bressler, S. L., Coppola, R. & Nakamura, R. Episodic multiregional cortical coherence at multiple frequencies during visual task performance. *Nature* **366**, 153–156 (1993).
7. Palva, J. M., Palva, S. & Kaila, K. Phase synchrony among neuronal oscillations in the human cortex. *J. Neurosci.* **25**, 3962–3972 (2005).
8. Wallis, J. D. & Miller, E. K. From rule to response: neuronal processes in the premotor and prefrontal cortex. *J. Neurophysiol.* **90**, 1790–1806 (2003).
9. Stokes, M. G. et al. Dynamic coding for cognitive control in prefrontal cortex. *Neuron* **78**, 364–375 (2013).
10. Mansouri, F. A., Freedman, D. J. & Buckley, M. J. Emergence of abstract rules in the primate brain. *Nat. Rev. Neurosci.* **21**, 595–610 (2020).
11. Voloh, B., Valiante, T. A., Everling, S. & Womelsdorf, T. Theta–gamma coordination between anterior cingulate and prefrontal cortex indexes correct attention shifts. *Proc. Natl Acad. Sci. USA* **112**, 8457–8462 (2015).
12. Fries, P. Rhythms for cognition: communication through coherence. *Neuron* **88**, 220–235 (2015).
13. Klimesch, W., Schack, B. & Sauseng, P. The functional significance of theta and upper alpha oscillations. *Exp. Psychol.* **52**, 99–108 (2005).

14. Cooper, P. S. et al. Frontal theta predicts specific cognitive control-induced behavioural changes beyond general reaction time slowing. *NeuroImage* **189**, 130–140 (2019).
15. Nigbur, R., Cohen, M. X., Ridderinkhof, K. R. & Stürmer, B. Theta dynamics reveal domain-specific control over stimulus and response conflict. *J. Cogn. Neurosci.* **24**, 1264–1274 (2011).
16. Donoghue, T. et al. Parameterizing neural power spectra into periodic and aperiodic components. *Nat. Neurosci.* **23**, 1655–1665 (2020).
17. Desimone, R. & Duncan, J. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**, 193–222 (1995).
18. Usher, M. & McClelland, J. L. The time course of perceptual choice: the leaky, competing accumulator model. *Psychol. Rev.* **108**, 550–592 (2001).
19. Verguts, T. Binding by random bursts: a computational model of cognitive control. *J. Cogn. Neurosci.* **29**, 1103–1118 (2017).
20. Senoussi, M., Moreland, J. C., Busch, N. A. & Dugué, L. Attention explores space periodically at the theta frequency. *J. Vis.* **19**, 22–22 (2019).
21. Kienitz, R., Schmid, M. C. & Dugué, L. Rhythmic sampling revisited: experimental paradigms and neural mechanisms. *Eur. J. Neurosci.* Advance online publication https://doi.org/10.1111/ejn.15489 (2021).
22. De Loof, E. et al. Preparing for hard times: scalp and intracranial physiological signatures of proactive cognitive control. *Psychophysiology* **56**, e13417 (2019).
23. Cavanagh, J. F., Cohen, M. X. & Allen, J. J. B. Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *J. Neurosci.* **29**, 98–105 (2009).
24. Luu, P., Tucker, D. M. & Makeig, S. Frontal midline theta and the error-related negativity: neurophysiological mechanisms of action regulation. *Clin. Neurophysiol.* **115**, 1821–1835 (2004).
25. Senoussi, M. et al. Pre-stimulus antero-posterior EEG connectivity predicts performance in a UAV monitoring task. In *Proceedings of 2016 International Conference on Systems, Man, and Cybernetics* (Canada): IEEE SMC, 1167–1172 (2017).
26. Kaiser, J. & Schütz-Bosbach, S. Proactive control without midfrontal control signals? The role of midfrontal oscillations in preparatory conflict adjustments. *Biol. Psychol.* **148**, 107747 (2019).
27. Nelli, S., Itthipuripat, S., Srinivasan, R. & Serences, J. T. Fluctuations in instantaneous frequency predict alpha amplitude during visual perception. *Nat. Commun.* **8**, 1–12 (2017).
28. Lopes da Silva, F. H., Vos, J. E., Mooibroek, J. & van Rotterdam, A. Relative contributions of intracortical and thalamo-cortical processes in the generation of alpha rhythms, revealed by partial coherence analysis. *Electroencephalogr. Clin. Neurophysiol.* **50**, 449–456 (1980).
29. Wutz, A., Melcher, D. & Samaha, J. Frequency modulation of neural oscillations according to visual task demands. *Proc. Natl Acad. Sci. USA* **115**, 1346–1351 (2018).
30. Samaha, J., Bauer, P., Cimaroli, S. & Postle, B. R. Top-down control of the phase of alpha-band oscillations as a mechanism for temporal prediction. *Proc. Natl Acad. Sci. USA* **112**, 8439–8444 (2015).
31. Lisman, J. E. & Jensen, O. The theta–gamma neural code. *Neuron* **77**, 1002–1016 (2013).
32. Siebenhühner, F., Wang, S. H., Palva, J. M. & Palva, S. Cross-frequency synchronization connects networks of fast and slow oscillations during visual working memory maintenance. *eLife* **5**, e13451 (2016).
33. Senoussi, M., Verbeke, P. & Verguts, T. Time-based binding as a solution to and a limitation for flexible cognition. *Front. Psychol.* **12**, 798061 (2022).
34. Axmacher, N. et al. Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proc. Natl Acad. Sci. USA* **107**, 3228–3233 (2010).
35. Kosciessa, J. Q., Grandy, T. H., Garrett, D. D. & Werkle-Bergner, M. Single-trial characterization of neural rhythms: potential and challenges. *NeuroImage* **206**, 116331 (2020).
36. Wolinski, N., Cooper, N. R., Sauseng, P. & Romei, V. The speed of parietal theta frequency drives visuospatial working memory capacity. *PLoS Biol.* **16**, e2005348 (2018).
37. Riddle, J., Scimeca, J. M., Cellier, D., Dhanani, S. & D'Esposito, M. Causal evidence for a role of theta and alpha oscillations in the control of working memory. *Curr. Biol.* **30**, 1748–1754 (2020).
38. Itthipuripat, S., Wessel, J. R. & Aron, A. R. Frontal theta is a signature of successful working memory manipulation. *Exp. Brain Res.* **224**, 255–262 (2013).
39. Mitchell, D. J., McNaughton, N., Flanagan, D. & Kirk, I. J. Frontal-midline theta from the perspective of hippocampal "theta". *Prog. Neurobiol.* **86**, 156–185 (2008).
40. Siapas, A. G., Lubenov, E. V. & Wilson, M. A. Prefrontal phase locking to hippocampal theta oscillations. *Neuron* **46**, 141–151 (2005).
41. Shenhav, A., Botvinick, M. M. & Cohen, J. D. The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron* **79**, 217–240 (2013).
42. Shenhav, A. et al. Toward a rational and mechanistic account of mental effort. *Annu. Rev. Neurosci.* **40**, 99–124 (2017).
43. Holroyd, C. B. & Yeung, N. Motivation of extended behaviors by anterior cingulate cortex. *Trends Cogn. Sci.* **16**, 122–128 (2012).
44. Pastötter, B., Dreisbach, G. & Bäuml, K.-H. T. Dynamic adjustments of cognitive control: oscillatory correlates of the conflict adaptation effect. *J. Cogn. Neurosci.* **25**, 2167–2178 (2013).
45. Verbeke, P. & Verguts, T. Neural synchrony for adaptive control. *J. Cogn. Neurosci.* **33**, 2394–2412 (2021).
46. Holroyd, C. B. & McClure, S. M. Hierarchical control over effortful behavior by rodent medial frontal cortex: a computational model. *Psychol. Rev.* **122**, 54–83 (2015).
47. Holroyd, C. B. & Verguts, T. The best laid plans: computational principles of anterior cingulate cortex. *Trends Cogn. Sci.* **25**, 316–329 (2021).
48. Holroyd, C. B., Ribas-Fernandes, J. J. F., Shahnazian, D., Silvetti, M. & Verguts, T. Human midcingulate cortex encodes distributed representations of task progress. *Proc. Natl Acad. Sci. USA* **115**, 6398–6403 (2018).
49. Womelsdorf, T., Johnston, K., Vinck, M. & Everling, S. Theta-activity in anterior cingulate cortex predicts task rules and their adjustments following errors. *Proc. Natl Acad. Sci. USA* **107**, 5248–5253 (2010).
50. Haynes, J.-D. et al. Reading hidden intentions in the human brain. *Curr. Biol.* **17**, 323–328 (2007).
51. Smith, E. H. et al. Widespread temporal coding of cognitive control in the human prefrontal cortex. *Nat. Neurosci.* **22**, 1883–1891 (2019).
52. Helfrich, R. F. et al. Neural mechanisms of sustained attention are rhythmic. *Neuron* **99**, 854–865 (2018).
53. Dugué, L., Roberts, M. & Carrasco, M. Attention reorients periodically. *Curr. Biol.* **26**, 1595–1601 (2016).
54. Dugué, L., McLelland, D., Lajous, M. & VanRullen, R. Attention searches nonuniformly in space and in time. *Proc. Natl Acad. Sci. USA* **112**, 15214–15219 (2015).
55. Fiebelkorn, I. C., Saalmann, Y. B. & Kastner, S. Rhythmic sampling within and between objects despite sustained attention at a cued location. *Curr. Biol.* **23**, 2553–2558 (2013).
56. Landau, A. N. & Fries, P. Attention samples stimuli rhythmically. *Curr. Biol.* **22**, 1000–1004 (2012).
57. Dugué, L., Xue, A. M. & Carrasco, M. Distinct perceptual rhythms for feature and conjunction searches. *J. Vis.* **17**(3), 22 (2017).
58. Holcombe, A. O. & Chen, W.-Y. Splitting attention reduces temporal resolution from 7 Hz for tracking one object to <3 Hz when tracking three. *J. Vis.* **13**(1), 12 (2013).
59. VanRullen, R. Perceptual cycles. *Trends Cogn. Sci.* **20**, 723–735 (2016).
60. Fiebelkorn, I. C. & Kastner, S. A rhythmic theory of attention. *Trends Cogn. Sci.* https://doi.org/10.1016/j.tics.2018.11.009 (2018).
61. Kienitz, R. et al. Theta rhythmic neuronal activity and reaction times arising from cortical receptive field interactions during distributed attention. *Curr. Biol.* **28**, 2377–2387 (2018).
62. Fiebelkorn, I. C. & Kastner, S. A rhythmic theory of attention. *Trends Cogn. Sci.* **23**, 87–101 (2019).
63. Helfrich, R. F. & Knight, R. T. Oscillatory dynamics of prefrontal cognitive control. *Trends Cogn. Sci.* **20**, 916–930 (2016).
64. de Vries, I. E. J., Slagter, H. A. & Olivers, C. N. L. Oscillatory control over representational states in working memory. *Trends Cogn. Sci.* **24**, 150–162 (2020).
65. Riddle, J., Vogelsang, D. A., Hwang, K., Cellier, D. & D'Esposito, M. Distinct oscillatory dynamics underlie different components of hierarchical cognitive control. *J. Neurosci.* **40**, 4945–4953 (2020).
66. Formica, S., González-García, C., Senoussi, M. & Brass, M. Neural oscillations track the maintenance and proceduralization of novel instructions. *NeuroImage* **232**, 117870 (2021).
67. Formica, S., González-García, C., Senoussi, M., Marinazzo, D. & Brass, M. Theta-phase connectivity between medial prefrontal and posterior areas underlies novel instructions implementation. Preprint at *bioRxiv* https://doi.org/10.1101/2022.02.23.481594 (2022).
68. Cavanagh, J. F., Zambrano-Vazquez, L. & Allen, J. J. B. Theta lingua franca: a common mid-frontal substrate for action monitoring processes. *Psychophysiology* **49**, 220–238 (2012).
69. Cohen, M. X. Midfrontal theta tracks action monitoring over multiple interactive time scales. *NeuroImage* **141**, 262–272 (2016).
70. Nee, D. E. & D'Esposito, M. The hierarchical organization of the lateral prefrontal cortex. *eLife* **5**, e12112 (2016).
71. Voytek, B. et al. Oscillatory dynamics coordinating human frontal networks in support of goal maintenance. *Nat. Neurosci.* **18**, 1318–1324 (2015).
72. Badre, D. & D'Esposito, M. Is the rostro-caudal axis of the frontal lobe hierarchical? *Nat. Rev. Neurosci.* **10**, 659–669 (2009).
73. Shahnazian, D., Senoussi, M., Krebs, R. M., Verguts, T. & Holroyd, C. B. Neural representations of task context and temporal order during action sequence execution. *Top. Cogn. Sci.* Advance online publication https://doi.org/10.1111/tops.12533 (2021).

74. Balaguer, J., Spiers, H., Hassabis, D. & Summerfield, C. Neural mechanisms of hierarchical planning in a virtual subway network. *Neuron* **90**, 893–903 (2016).
75. Driel, J., van, Sligte, I. G., Linders, J., Elport, D. & Cohen, M. X. Frequency band-specific electrical brain stimulation modulates cognitive control processes. *PLoS ONE* **10**, e0138984 (2015).
76. Lehr, A., Henneberg, N., Nigam, T., Paulus, W. & Antal, A. Modulation of conflict processing by theta-range tACS over the dorsolateral prefrontal cortex. *Neural Plast.* **2019**, e6747049 (2019).
77. Riddle, J. & Frohlich, F. Targeting neural oscillations with transcranial alternating current stimulation. *Brain Res.* **1765**, 147491 (2021).
78. O'Connell, R. G., Dockree, P. M. & Kelly, S. P. A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nat. Neurosci.* **15**, 1729–1735 (2012).
79. Wyart, V., de Gardelle, V., Scholl, J. & Summerfield, C. Rhythmic fluctuations in evidence accumulation during decision making in the human brain. *Neuron* **76**, 847–858 (2012).
80. Hagemann, D., Hewig, J., Walter, C. & Naumann, E. Skull thickness and magnitude of EEG alpha activity. *Clin. Neurophysiol.* **119**, 1271–1280 (2008).
81. Voytek, B. et al. Hemicraniectomy: a new model for human electrophysiology with high spatio-temporal resolution. *J. Cogn. Neurosci.* **22**, 2491–2502 (2010).
82. Klimesch, W. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Rev.* **29**, 169–195 (1999).
83. Mierau, A., Klimesch, W. & Lefebvre, J. State-dependent alpha peak frequency shifts: experimental evidence, potential mechanisms and functional implications. *Neuroscience* **360**, 146–154 (2017).
84. Haegens, S., Cousijn, H., Wallis, G., Harrison, P. J. & Nobre, A. C. Inter- and intra-individual variability in alpha peak frequency. *NeuroImage* **92**, 46–55 (2014).
85. Minami, S., Oishi, H., Takemura, H. & Amano, K. Inter-individual differences in occipital alpha oscillations correlate with white matter tissue properties of the optic radiation. *eNeuro* **7**, 2 (2020).
86. Grandy, T. H. et al. Peak individual alpha frequency qualifies as a stable neurophysiological trait marker in healthy younger and older adults. *Psychophysiology* **50**, 570–582 (2013).
87. Jafari, Z., Kolb, B. E. & Mohajerani, M. H. Neural oscillations and brain stimulation in Alzheimer's disease. *Prog. Neurobiol.* **194**, 101878 (2020).
88. Pistono, A. et al. Language network connectivity increases in early Alzheimer's disease. *J. Alzheimers Dis.* **82**, 447–460 (2021).
89. Brem, A.-K. & Sensi, S. L. Towards combinatorial approaches for preserving cognitive fitness in aging. *Trends Neurosci.* **41**, 885–897 (2018).
90. Chen, L., Chung, S. W., Hoy, K. E. & Fitzgerald, P. B. Is theta burst stimulation ready as a clinical treatment for depression? *Expert Rev. Neurother.* **19**, 1089–1102 (2019).
91. Slobodskoy-Plusnin, J. Behavioral and brain oscillatory correlates of affective processing in subclinical depression. *J. Clin. Exp. Neuropsychol.* **40**, 437–448 (2018).
92. Cohen, M. X. A neural microcircuit for cognitive conflict detection and signaling. *Trends Neurosci.* **37**, 480–490 (2014).
93. Silvetti, M., Vassena, E., Abrahamse, E. & Verguts, T. Dorsal anterior cingulate–brainstem ensemble as a reinforcement meta-learner. *PLoS Comput. Biol.* **14**, e1006370 (2018).
94. Sara, S. J. Locus Coeruleus in time with the making of memories. *Curr. Opin. Neurobiol.* **35**, 87–94 (2015).
95. Silvetti, M., Wiersema, J. R., Sonuga-Barke, E. & Verguts, T. Deficient reinforcement learning in medial frontal cortex as a model of dopamine-related motivational deficits in ADHD. *Neural Netw.* **46**, 199–209 (2013).
96. Bonnefond, M., Kastner, S. & Jensen, O. Communication between brain areas based on nested oscillations. *eNeuro* **4**, 2 (2017).
97. Verbeke, P. & Verguts, T. Learning to synchronize: how biological agents can couple neural task modules for dealing with the stability–plasticity dilemma. *PLoS Comput. Biol.* **15**, e1006604 (2019).
98. Bartos, M., Vida, I. & Jonas, P. Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks. *Nat. Rev. Neurosci.* **8**, 45–56 (2007).
99. Whittington, M. A., Cunningham, M. O., LeBeau, F. E. N., Racca, C. & Traub, R. D. Multiple origins of the cortical gamma rhythm. *Dev. Neurobiol.* **71**, 92–106 (2011).
100. Tiesinga, P. & Sejnowski, T. J. Cortical enlightenment: are attentional gamma oscillations driven by ING or PING? *Neuron* **63**, 727–732 (2009).
101. Wang, X. J., Golomb, D. & Rinzel, J. Emergent spindle oscillations and intermittent burst firing in a thalamic model: specific neuronal mechanisms. *Proc. Natl Acad. Sci. USA* **92**, 5577–5581 (1995).
102. Gips, B., Eerden, J. P. J. Mvander & Jensen, O. A biologically plausible mechanism for neuronal coding organized by the phase of alpha oscillations. *Eur. J. Neurosci.* **44**, 2147–2161 (2016).
103. Wong, K.-F. & Wang, X.-J. A recurrent network mechanism of time integration in perceptual decisions. *J. Neurosci.* **26**, 1314–1328 (2006).
104. Jia, X., Xing, D. & Kohn, A. No consistent relationship between gamma power and peak frequency in macaque primary visual cortex. *J. Neurosci.* **33**, 17–25 (2013).
105. Siegel, M., Warden, M. R. & Miller, E. K. Phase-dependent neuronal coding of objects in short-term memory. *Proc. Natl Acad. Sci. USA* **106**, 21341–21346 (2009).
106. Atallah, B. V. & Scanziani, M. Instantaneous modulation of gamma oscillation frequency by balancing excitation with inhibition. *Neuron* **62**, 566–577 (2009).
107. Palestro, J. J., Weichart, E., Sederberg, P. B. & Turner, B. M. Some task demands induce collapsing bounds: evidence from a behavioral analysis. *Psychon. Bull. Rev.* **25**, 1225–1248 (2018).
108. Botvinick, M. M. & Cohen, J. D. The computational and neural basis of cognitive control: charted territory and new frontiers. *Cogn. Sci.* **38**, 1249–1285 (2014).
109. Müller, M. G., Papadimitriou, C. H., Maass, W. & Legenstein, R. A model for structured information representation in neural networks. *eNeuro* **7**, 3 (2020).
110. Peirce, J. et al. PsychoPy2: experiments in behavior made easy. *Behav. Res. Methods* **51**, 195–203 (2019).
111. Dalmaijer, E. S., Mathôt, S. & Van der Stigchel, S. PyGaze: an open-source, cross-platform toolbox for minimal-effort programming of eyetracking experiments. *Behav. Res. Methods* **46**, 913–921 (2014).
112. Gramfort, A. et al. MEG and EEG data analysis with MNE-Python. *Front. Neurosci.* **7**, 267 (2013).
113. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
114. Manning, J. R., Jacobs, J., Fried, I. & Kahana, M. J. Broadband shifts in local field potential power spectra are correlated with single-neuron spiking in humans. *J. Neurosci.* **29**, 13613–13620 (2009).
115. He, B. J. Scale-free brain activity: past, present, and future. *Trends Cogn. Sci.* **18**, 480–487 (2014).
116. Voytek, B. et al. Age-related changes in 1/*f* neural electrophysiological noise. *J. Neurosci.* **35**, 13257–13265 (2015).
117. Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007).
118. Gerster, M. et al. Separating neural oscillations from aperiodic 1/*f* activity: challenges and recommendations. Preprint at https://www.biorxiv.org/content/10.1101/2021.10.15.464483v1 (2021).
119. Kerby, D. S. The simple difference formula: an approach to teaching nonparametric correlation. *Compr. Psychol.* **3**, 11.IT.3.1 (2014).
120. Vallat, R. Pingouin: statistics in Python. *J. Open Source Softw.* **3**, 1026 (2018).
121. Rousselet, G. A. & Pernet, C. R. Improving standards in brain–behavior correlation analyses. *Front. Hum. Neurosci.* **6**, 119 (2012).

Acknowledgements

The authors thank C. Buc Calderon for fruitful discussions and comments on the manuscript. M.S. and T.V. were supported by grant G012816 from Research Foundation Flanders. M.S., T.V. and E.D.L. were supported by grant BOF17-GOA-004 from the Research Council of Ghent University. P.V. was supported by grant 1102519N from Research Foundation Flanders. K.D. was supported by FWO [PEGASUS]^2 Marie Skłodowska-Curie fellowship 12T9717N. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

M.S., D.T. and T.V. designed the study. M.S., P.V. and T.V. developed the model. M.S. and E.D.L. collected the data. M.S. analysed model simulations, and behavioural and EEG data. M.S. and T.V. wrote the manuscript. All of the authors discussed the results and commented on the manuscript.

Competing interests

The authors declare no competing interests.

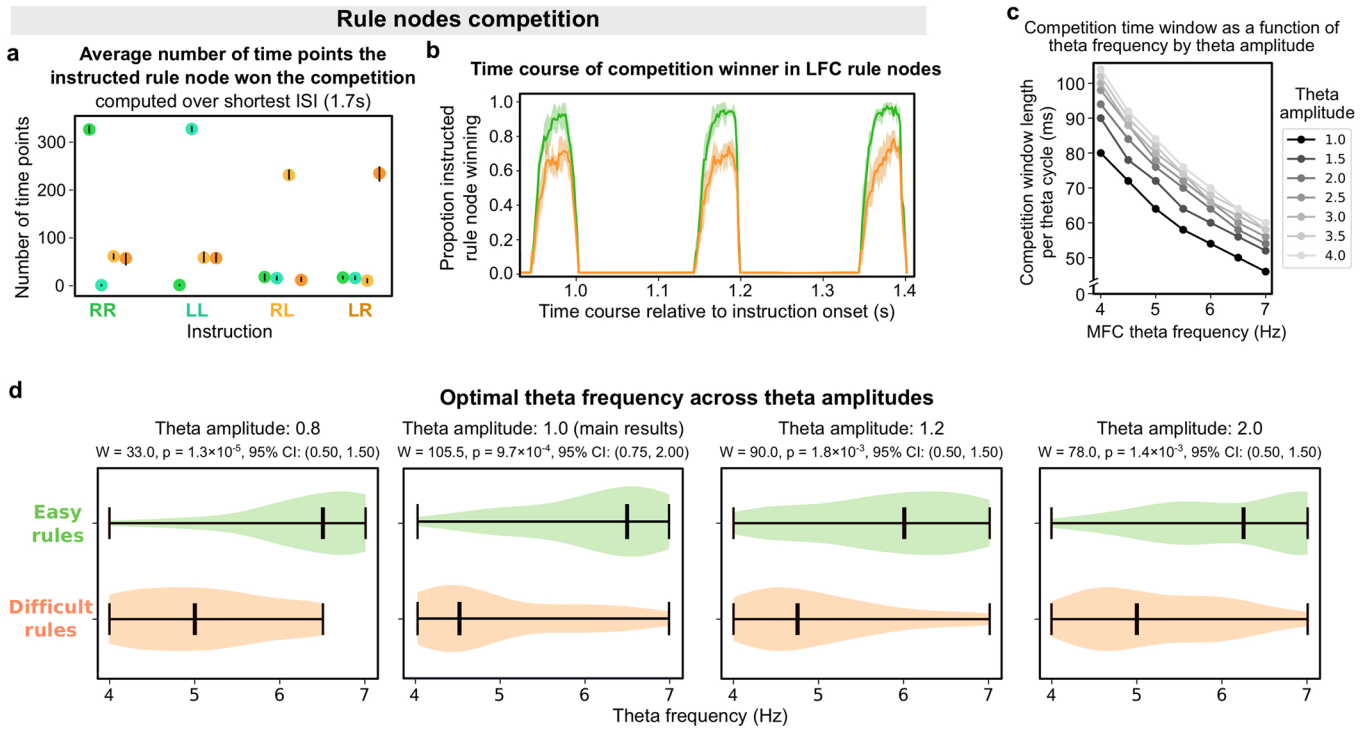
Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41562-022-01335-5.

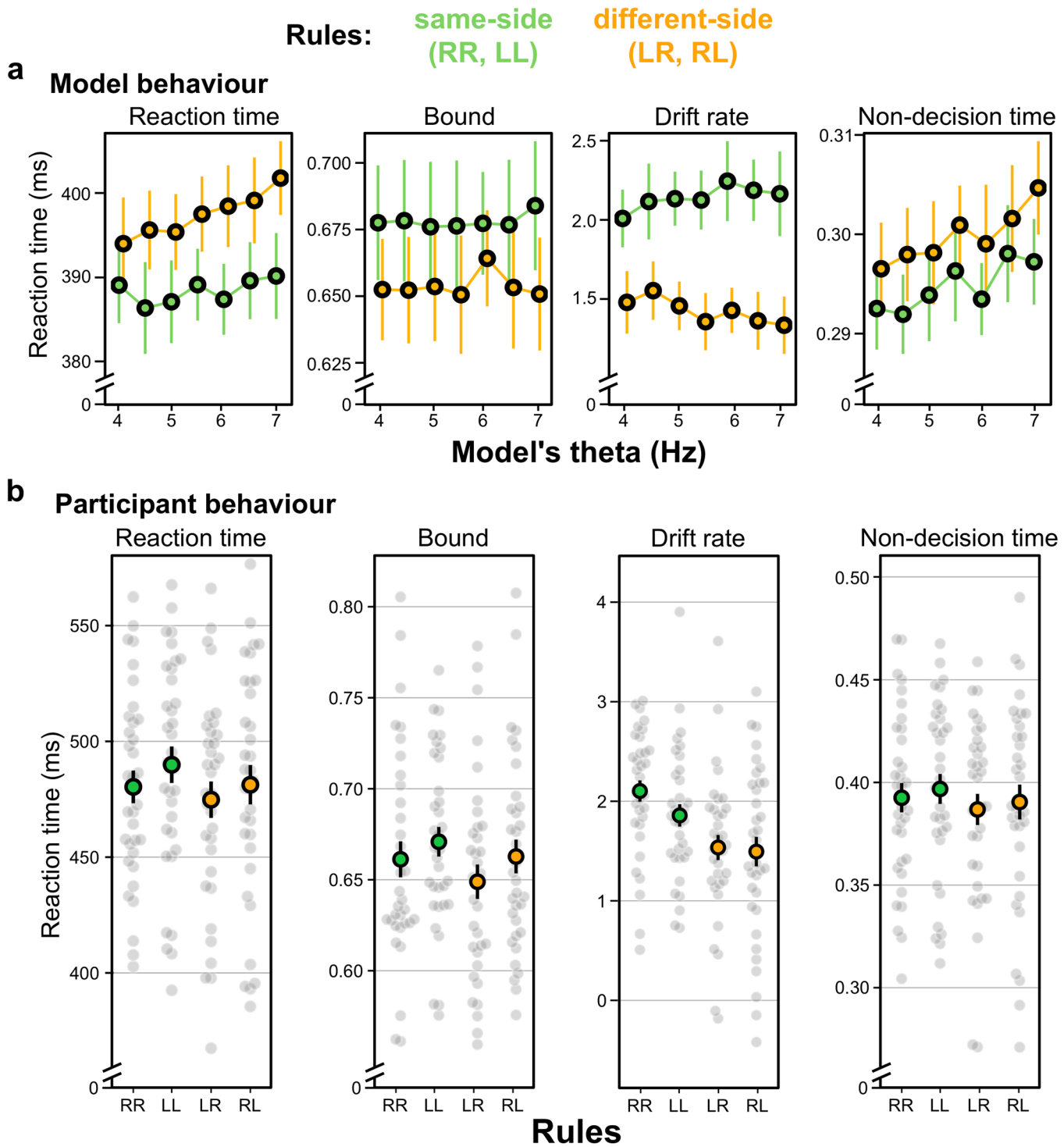
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41562-022-01335-5.

Correspondence and requests for materials should be addressed to Mehdi Senoussi.

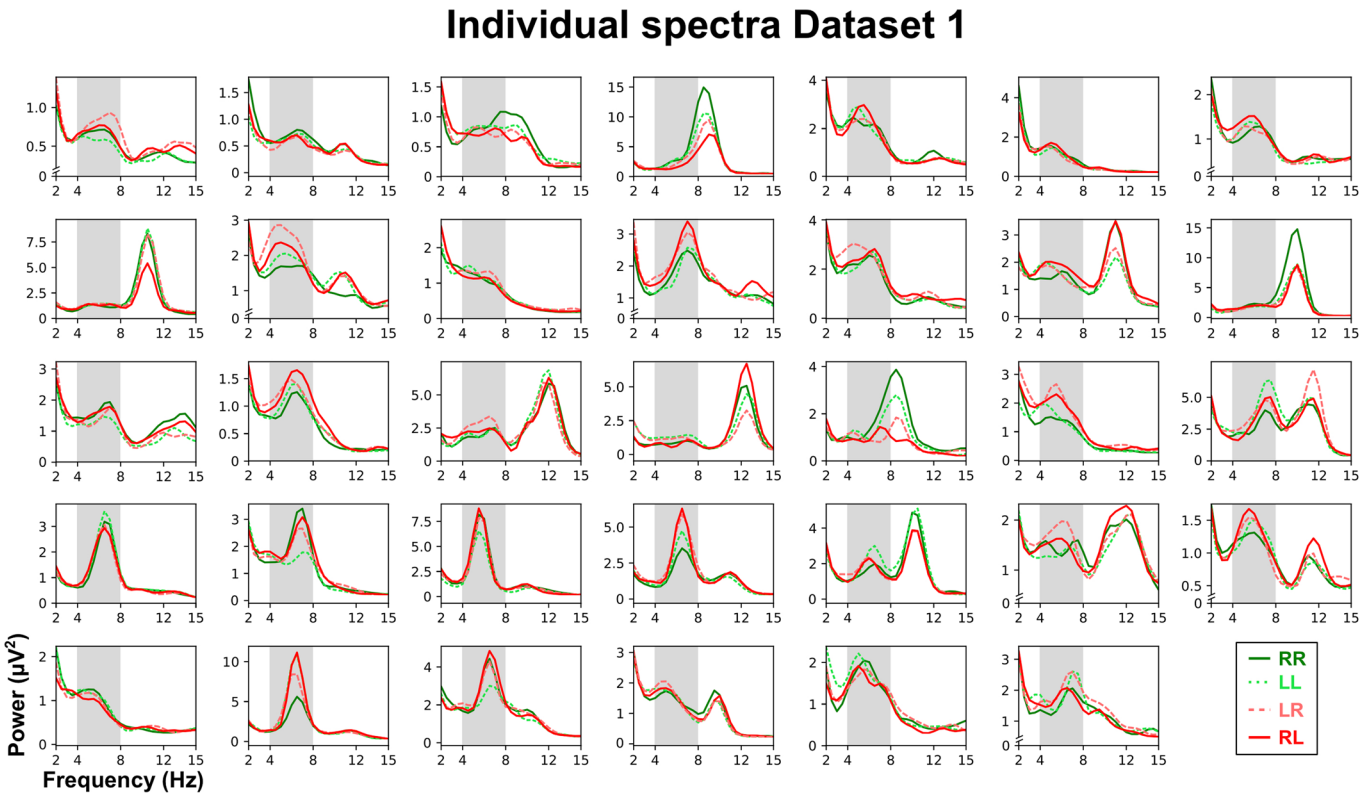
Peer review information *Nature Human Behaviour* thanks James Cavanagh, Sirawaj Itthipuripat and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.



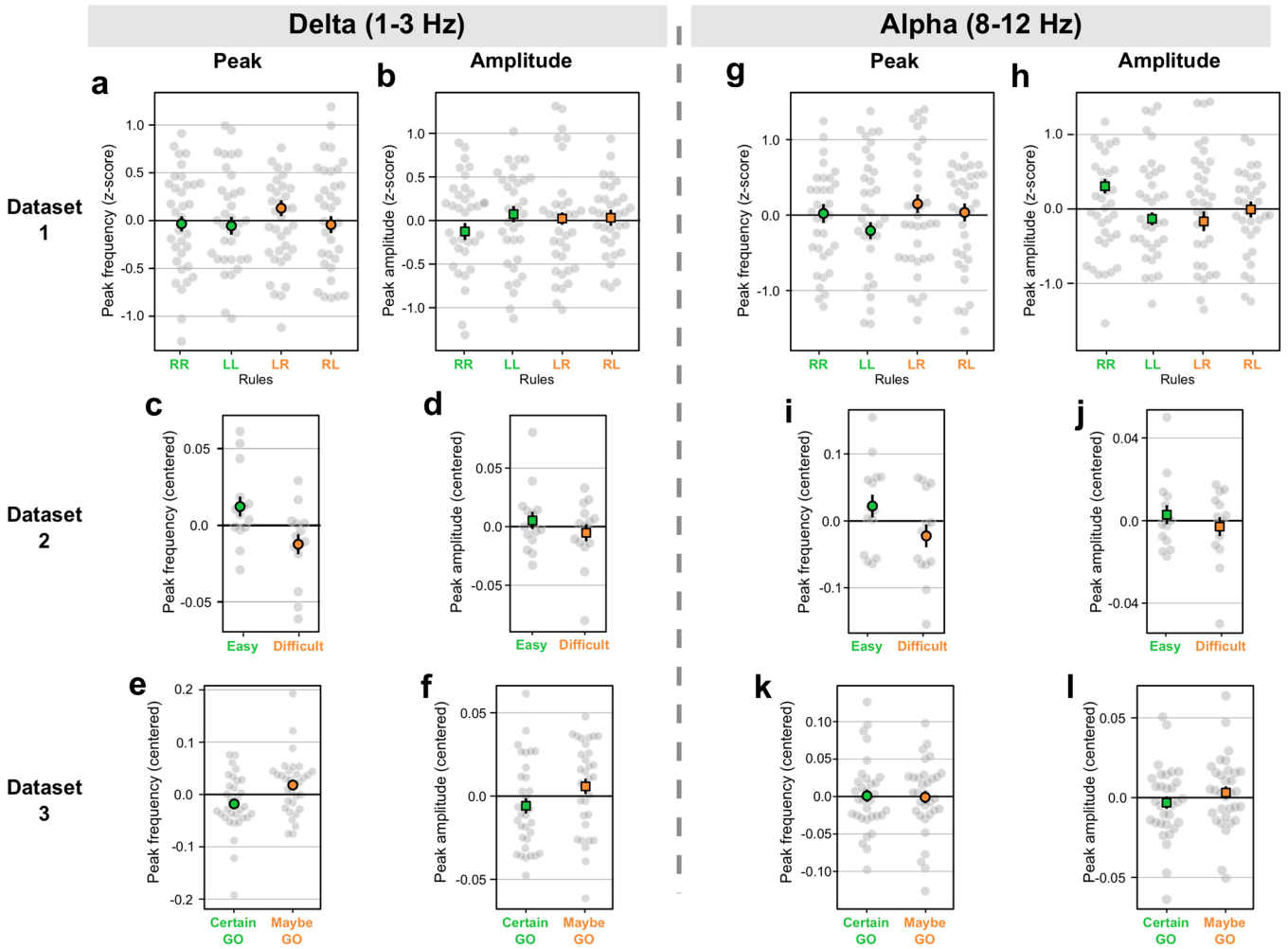
Extended Data Fig. 1 | Rule node competition and effect of MFC theta amplitude. a Number of time points the correct rule node won the competition in the LFC unit depending on the instructed rule. Black bars represent the 95% Confidence Interval (smaller than dot size). The colour of each dot represents the rule, as represented in the rule names on the x-axis. b Time course of proportion of correct rule node winning the competition across a sample of the ISD. Only two rule nodes are shown for clarity, one easy (green curve) and one difficult (orange curve). Shaded area represents the 95% Confidence Interval. c Effect of MFC theta amplitude and frequency on competition time window. Average competition window length, in milliseconds, for one theta cycle, that is one competition window, as a function of theta frequency for different theta amplitude. Varying the amplitude (the different lines) shows that although the competition window increases with amplitude, this effect reaches a ceiling around amplitude values of 2-3. Each line represents simulations at different levels of MFC theta frequency with a fixed theta amplitude. Line color represents MFC theta amplitude as represented in the legend. d Varying MFC theta amplitude from 0.8 to 2.0 yielded similar results concerning what was the optimal theta frequency depending on rule difficulty ($n = 34$ simulations per theta frequency and for each theta amplitude, two-sided Wilcoxon sign-rank test: all $W > 33$, all $ps < 0.002$). Data are presented as violin plots, left- and right-most bars represent extrema, middle bar represent the median. Distribution density is represented by violin plot width.



Extended Data Fig. 2 | Reaction times and control analyses with DDM. Reaction times and DDM parameters (bound, drift rate and non-decision time) estimated in model and participant data with the EZ-Diffusion model (see Methods). a Reaction time and DDM parameters of model performance by rule difficulty (same-side, different-side) and theta frequency (4 to 7Hz, steps of 0.5Hz), n = 34 simulations per theta frequency. We ran a repeated-measure 2x7 ANOVA with factors rule difficulty (2 levels) and theta frequency (7 levels). There was a main effect of rule difficulty in all measures, that is reaction times, bound, drift rate and non-decision time, (all Fs(1, 33) > 98.96, all ps < 10-10, all $\eta^2 > 0.20$). There was a main effect of theta frequency in reaction times, drift rate and nondecision time (all Fs(6, 33) > 3.20, all ps < 0.006, all η 2 > 0.01). There was a significant rule-difficulty – theta-frequency interaction in drift rate ($F(6, 33) = 6.69$, $p < 0.001$, $\eta^2 = 0.021$). Error bars represent standard deviation across simulations. b Reaction time and DDM parameters (bound, drift rate and nondecision time) estimated on participants' data grouped by rule (n = 34 participants). Data were collapsed across ISD to avoid data sparsity. We ran a repeated-measure 2x2 ANOVA with factors target-location (2 levels: Left, Right) and hand (2 levels: Left, Right). Only drift rate showed a significant interaction between hand and target-location ($F(1, 33) = 31.86$, $p < 0.001$, $\eta^2 = 0.21$), as well as a main effect of hand ($F(1, 33) = 6.65$, $p = 0.014$, $\eta^2 = 0.02$). Reaction time and bound showed a significant effect of hand (Reaction time: $F(1, 33) = 4.62$, $p = 0.039$, $\eta^2 = 0.04$; bound: $F(1, 33) = 4.84$, $p = 0.034$, $\eta^2 = 0.03$). All other effects were not significant (all Fs(1, 33) < 3.29, all ps > 0.079). Data are presented as mean values, error bars represent s.e.m. Smaller grey dots represent individual participants' data.



Extended Data Fig. 3 | Raw spectra of individual participants per rule in Dataset 1. The grey area represents the theta frequency band.



Extended Data Fig. 4 | Control analyses on the effect of peak frequency and amplitude of nearby frequency bands. Panels a to f: delta frequency band (1-3Hz). Panels g to l: alpha frequency band (8-12Hz). Two-way repeated-measure ANOVAs were used for all Dataset 1 data. Two-sided Wilcoxon sign-rank tests were used for all Dataset 2 and 3 data. a Peak frequency of delta band oscillations in Dataset 1 (all Fs < 0.78, all ps > 0.384). b Peak amplitude of delta band oscillations in Dataset 1 (all Fs < 0.76, all ps > 0.390). c Peak frequency of delta band oscillations in Dataset 2 (W = 26, p = 0.104, r = 0.50, 95% CI = (−0.00, 0.04)). d Peak amplitude of delta band oscillations in Dataset 2 (W = 49, p = 0.855, r = 0.07, 95% CI = (−0.02, 0.03)). e Peak frequency of delta band oscillations in Dataset 3 (W = 173, p = 0.091, r = − 0.34, 95% CI = (−0.07, 0.01)). f Peak amplitude of delta band oscillations in Dataset 3 (W = 195, p = 0.200, r = −0.26, 95% CI = (−0.04, 0.01)). g Peak frequency of alpha band oscillations in Dataset 1 (all Fs < 1.71, all ps > 0.199). h Peak amplitude of alpha band oscillations in Dataset 1. There was a main effect of target-location ($F(1, 33) = 5.32$, $p = 0.027$, $\eta 2 = 0.051$; uncorrected for multiple comparisons). i Peak frequency of alpha band oscillations in Dataset 2 (W = 29, p = 0.153, r = 0.44, 95% CI = (−0.05, 0.13)). j Peak amplitude of alpha band oscillations in Dataset 2 (W = 51, p = 0.951, r = 0.03, 95% CI = (−0.02, 0.02)). k Peak frequency of alpha band oscillations in Dataset 3 (W = 244, p = 0.715, r = − 0.07, 95% CI = (−0.05, 0.03)). l Peak amplitude of alpha band oscillations in Dataset 3 (W = 218, p = 0.394, r = −0.17, 95% CI = (−0.03, 0.01)). Data are presented as mean values, error bars represent s.e.m. computed over n = 34, 14 and 33 participants for Dataset 1, 2 and 3, respectively. Smaller grey dots represent individual participants' data.

nature research

Corresponding author(s): Mehdi Senoussi (mehdi.senoussi@ugent.be)

Last updated by author(s): Feb 24, 2022

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	iView RED250mobile software for eye-tracking recording v1.0, BrainVision Pycorder v1.0 for EEG recording, Python 2.7 32bit, Psychopy toolbox v.3.0 for stimuli presentation, Pygaze toolbox for communication between stimulus computer and eyetracker computer version 0.7.1.
Data analysis	Python 3.7 64bit, Numpy Python library version 1.16.5, Scipy Python library version 1.3.1, MNE Python toolbox version 0.21, FOOOF toolbox version 1.0.0, StatsModels Python package version 0.10.1, Pingouin Python package v0.5.0. Custom code was written for the model and for behavioral and EEG analysis, all scripts are available on the Github repository of the project: https://github.com/mehdisenoussi/theta_shift_cog_control.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Raw behavioral, eye-tracking and EEG data can be found at this Open Science Framework repository: https://osf.io/nwh87/?view_only=b11ee1f860804da582c816fe8acdecad

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For Dataset 1, sample size was not computed a priori: we aimed for more than 30 participants, which is on the upper bound of current EEG practices. First, we recruited 35 participants to reach a total number of more than 30 participants after drop-out, considering a ~10% drop-out rate due to noise-corrupted data or other issues related to participants' task performance. The sample size after exclusions dropped at 30 participants, we thus tested 4 more participants, which were all included, bringing the sample size to 34 participants. Assuming a medium effect size and aiming for a power of 0.8 in a within-subject repeated measures ANOVA analysis, the study would require a sample of 32 participants.
Data exclusions	For Dataset 1, five participants were excluded from the analysis: two completed less than 5 blocks, one had less than 200 trials after trial rejection based on eye-tracking data, one had poor overall behavioral performances (i.e. less than 50% overall accuracy), and one participant was left handed. For Datasets 2 and 3, participants were excluded from the analysis if no peak in the theta frequency range could be found using the FOOOF toolbox at electrode FCz.
Replication	Multiple model simulations allowed to verify the reproducibility of model results. Replications of the EEG findings of Dataset 1 (i.e. lower frequency of frontal mid-line theta in difficult than in easy conditions) was performed using two independent datasets (see manuscript).
Randomization	Trial types (e.g. trial instructions) were randomized independently for each participant for all three datasets. All datasets were repeated-measures design there was thus no randomization of participants.
Blinding	All three datasets were repeated-measures designs, there was therefore no participant groups to conduct a blinded procedure.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	For Dataset 1, thirty-nine human participants were recruited for this experiment (M ± STD = 23.7 ± 4.5 years old, range: 18-41 years old; 27 females). All participants had normal or corrected-to-normal vision and no history of neurological problems.
Recruitment	For Dataset 1, participants were recruited through the online participant management system SONA, contracted by Ghent University. Most participants were students of Ghent University. We do not believe that this had an impact on the results.
Ethics oversight	For Dataset 1, all participants provided written informed consent and received monetary compensation for their participation. The experiment was approved by the local ethics committee (Faculty of Psychology and Educational Sciences, Ghent University).

Note that full information on the approval of the study protocol must also be provided in the manuscript.