Causal evidence for a domain-specific role of left superior frontal sulcus in human perceptual decision-making

Miguel Barretto García,[1†] Marcus Grueschow,[1†] Marius Moisa,[1] Rafael Polania,[2] and Christian C. Ruff[1]

[†] These authors contributed equally to this paper.

Affiliations:

[1] Zurich Center for Neuroeconomics (ZNE), Department of Economics, University of Zurich, Zurich, Switzerland.

[2] Decision Neuroscience Lab, Dept. of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland.

ABSTRACT

Humans and animals can flexibly choose their actions based on different information, ranging from objective states of the environment (e.g., apples are bigger than cherries) to subjective preferences (e.g., cherries are tastier than apples). Whether the brain instantiates these different choices by recruiting either specialized or shared neural circuitry remains debated. Specifically, domain-general theories of prefrontal cortex (PFC) function propose that prefrontal areas flexibly process either perceptual or value-based evidence depending on what is required for the present choice, whereas domain-specific theories posit that PFC sub-areas, such as the left superior frontal sulcus (SFS), selectively integrate evidence relevant for perceptual decisions. Here we comprehensively test the functional role of the left SFS for choices based on perceptual and value-based evidence, by combining fMRI with a behavioural paradigm, computational modelling, and transcranial magnetic stimulation. Confirming predictions by a sequential sampling model, we show that TMS-induced excitability reduction of the left SFS selectively changes the processing of decision-relevant perceptual information and associated neural processes. In contrast, value-based decision making and associated neural processes remain unaffected. This specificity of SFS function is evident at all levels of analysis (behavioural, computational, and neural, including functional connectivity), demonstrating that the left SFS causally contributes to evidence integration for perceptual but not value-based decisions.

INTRODUCTION

Humans and animals choose their actions based on diverse types of information. For instance, during perceptual decisions, organisms discriminate choice options based on objective states of the environment (e.g., apples are bigger than cherries), whereas during value-based decisions, organisms make choices based on subjective preferences (e.g., cherries are tastier than apples). The degree to which the brain instantiates these different types of choices by specialized or shared neural circuitry is debated (Summerfield and Tsetsos 2012, Tajima et al, 2016, Polania et al 2015, Deco et al., 2013).

From a computational perspective, sequential-sampling frameworks propose that all types of choices should involve the continuous accumulation of evidence about the choice alternatives until a decision criterion is met and a motor action is executed (Busemeyer and Townsend 1993, Dutilh & Rieskamp 2016, Gold & Shadlen 2007, Krajbich 2019, Mazurek et al 2003, Ratcliff and Mckoon 1988, Usher & McClelland 2001). Computational modelling studies have largely confirmed that such a framework is applicable to both perceptual (Mazurek et al 2003, Moran 2015, Ratcliff and Mckoon 1988, Ratcliff & Rouder 1998) as well as value-based decisions (Busemeyer & Townsend 1993, Tajima et al 2016, Usher & McClelland 2001), suggesting that evidence accumulation processes may constitute a domain-general decision mechanism (Bogacz et al 2006, Dutilh & Rieskamp 2016, Gerstner et al 2012, Smith & Ratcliff 2004, Usher & McClelland 2001).

Empirical support for the neural instantiation of such domain-general choice mechanisms, however, is sparse. Neuroimaging studies in the brains of rodents, monkeys and humans have largely focussed on perceptual evidence accumulation (Brody and Hanks 2016, Gold and Shadlen 2007, Hanks et al 2015, Mulder et al 2014). Seminal human imaging studies of perceptual decision making have repeatedly implicated the superior frontal sulcus (SFS), a portion of the dorsolateral prefrontal cortex (DLPFC), in that process (Heekeren et al 2004, Heekeren et al 2008, Kayser et al 2010, Keuken et al 2014, Mulder et al 2014, Ploran et al

2007, White et al 2012). Moreover, disruption of the human left SFS with non-invasive brain stimulation impacted on behavioural performance and response speed in a face-house classification task, in a manner consistent with reduction of evidence accumulation processes as defined in sequential sampling models (Philiastides et al 2011). This has been taken as causal evidence that the left SFS supports perceptual evidence accumulation during decision making.

However, this specificity has been questioned, since the SFS is part of a frontoparietal network that has long been associated with domain-general attentional processing and cognitive control (Dodds et al 2011, Marek and Dosenbach 2018, Scolari et al 2015, Dixon et al 2018). It is therefore often argued that behavioural changes following SFS disruption may not reflect specific decision-making deficits but rather reductions in general levels of cognitive performance (Muhle-Karbe et al 2018, Mayer et al., 2017, Grueschow et al., 2020). Relatedly, animal lesion studies have questioned whether the structures showing activity resembling integration processes really serve this function (Erlich et al 2015, Hanks et al 2015, Piet et al 2017). For instance, while earlier work in rodents associated neural firing with an accumulation mechanism in the DLPFC during the perceptual choice process (Ding and Gold 2012b, Heitz and Schall 2012, Kim and Shadlen 1999, Mante et al 2013, Purcell et al 2010), recent inactivation studies argue that DLPFC activity appears more consistent with categorical encoding of decision choices, potentially even resembling a post-categorization memory trace (Erlich et al 2015, Hanks et al 2015, Piet et al 2017). Hence, while some causal contribution of SFS to perceptual decision making is undisputed, its precise contribution to the decision-making processes, and the selectivity of this contribution, remains to be resolved, in particular in relation to processes predicted by sequential sampling models.

The ambiguity about the SFS's precise causal function is also apparent in contradictory findings of human non-invasive brain stimulation studies employing the DDM (Philiastides et al 2011, Rahnev et al 2016). That is, one study reported that SFS disruption

during a speeded perceptual categorization task reduced accuracy and increased response times (Philiastides et al 2011) and found associated decreases in drift rate, the DDM parameter describing the efficiency of sensory evidence integration. In contrast, another human brain stimulation study suggested that behavioural changes due to SFS disruption during a perceptual 2AFC task reflect decreased response caution, characterized by faster response speed but decreased choice precision. Simulations with the same DDM modelling framework (Rahnev et al 2016) suggested that the decision threshold parameter, rather than the drift-rate, could account for individual behavioral changes. Simultaneously acquired fMRI data suggested that SFS does not code the rate of integration but rather the necessary amount of evidence to be accumulated for the perceptual choice at hand (Rahnev et al 2016). These inconsistent findings reveal a substantial level of uncertainty in the literature about which precise functional role the SFS serves during perceptual decision making.

Last but not least, the domain-specificity of the SFS contribution is unclear, since direct comparisons with other types of decision-making are largely missing. Even though some previous studies have suggested that DLPFC activity may reflect value-based evidence integration (Basten et al 2010, Sokol-Hessner et al 2012), it is hard to directly compare the implicated neural processes to those underlying perceptual choices, due to major differences in the stimuli and experimental approaches classically used in each domain (Balleine 2007, Gold and Shadlen 2007, Heekeren et al 2004, Krawczyk 2002). While there are good theoretical reasons to believe that common mechanisms may underlie both perceptual and value-based choices (Chawla and Miyapuram 2018, Summerfield and Tsetsos 2012), the number of studies directly comparing the neural mechanisms between both choice domains is surprisingly limited. The few existing studies have all used correlational neuroimaging techniques (Grueschow et al 2015, Polania et al 2014), and no study to date has used causal brain-stimulation techniques in combination with neuroimaging to directly compare the neuro-computational contributions of the SFS to both types of choice.

Here we provide such a detailed comparison. We applied continuous theta-burst transcranial magnetic stimulation (cTBS) followed by functional magnetic resonance imaging (fMRI) while human participants alternated between perceptual and value-based choices based on matched stimuli and involving the same motor responses. We modeled the observed behavioural changes with the DDM, allowing us to causally associate the stimulated SFS region to specific underlying latent sub-processes of the unfolding decision (Mulder et al 2014, Polania et al 2015). This computational framework provides us with clear testable hypotheses regarding possible effect patterns on the behavioural, computational, and neural levels. For instance, if SFS neurons indeed selectively accumulate perceptual evidence, we should find that their inhibition by cTBS leads to decreases in choice precision and increases in reaction times, a behavioural pattern that corresponds to a decrease in the DDM drift-rate parameter (Philiastides et al 2011) and to concurrent increases in BOLD signals (caused by prolonged neural evidence accumulation; **Fig 1a-c**). Critically, a different pattern can be expected when SFS neurons are involved in setting the criterion, i.e., determining the amount of evidence that needs to be accumulated for a perceptual choice to be taken. In this case, SFS inhibition should result in decreases in both choice precision and reaction times, a decrease in the DDM boundary parameter (Rahnev et al 2016), and a reduction in associated neural activity due to the lower amount of evidence accumulated during the shorter response time (**Fig. 1d-f**). Here we directly test these two contrasting scenarios, by characterizing the behavioral, neural, and neuro-computational consequences of cTBS to the left superior frontal sulcus (SFS). Crucially, we also investigate for both possible outcomes whether the functional contribution of the SFS during decision making is indeed specific for perceptual choices, by comparing the results between the two matched types of choices.

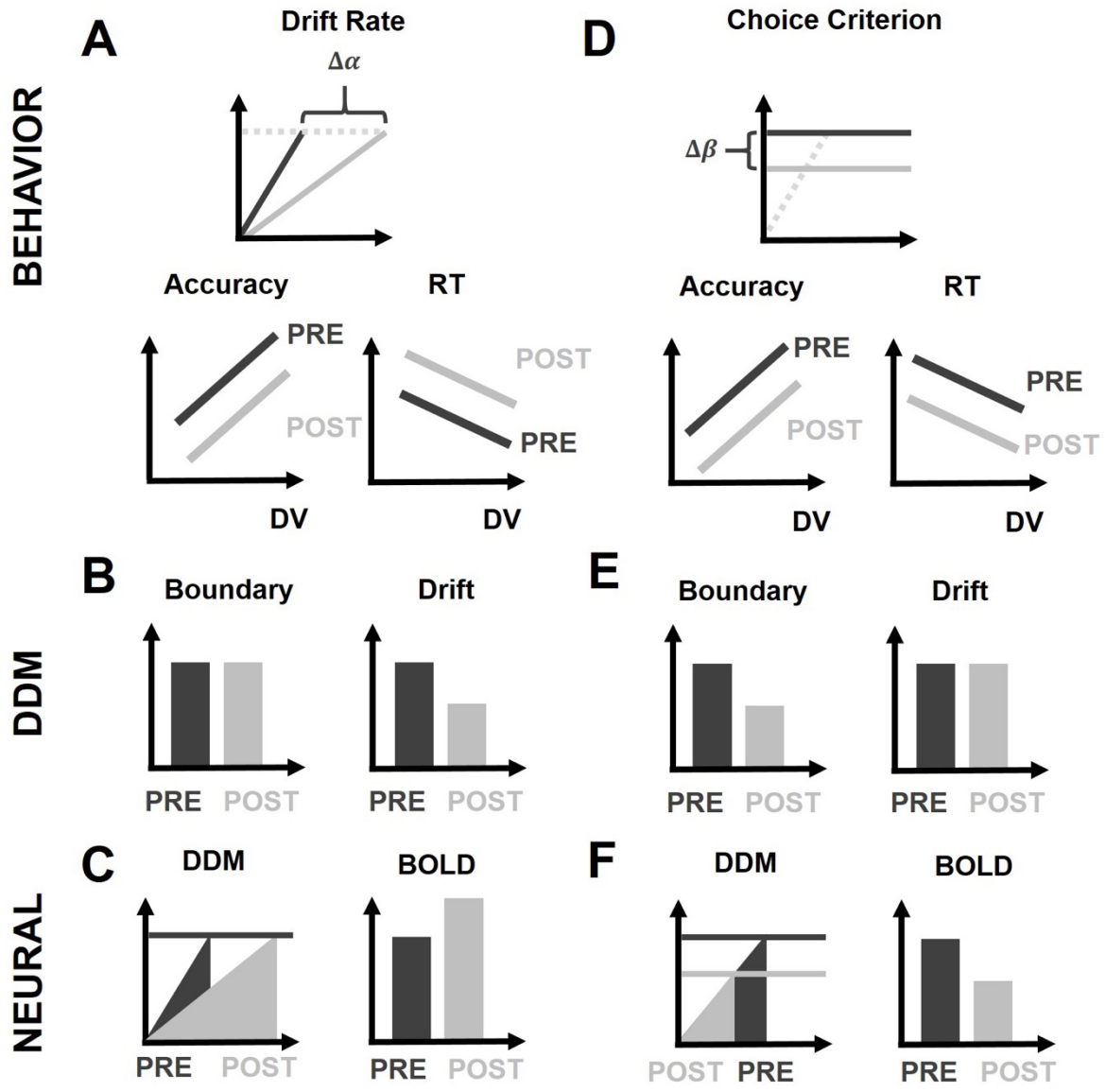

FIGURE 1. Study hypotheses. *Scenario 1: left SFS is causally involved in evidence accumulation.* Theta-burst induced inhibition of left SFS should lead to reduced evidence accumulation **(A)**, expressed as lower accuracy **(A, 2nd row, left)**, slowing of RTs **(A, 2nd row, right)**, and a reduction of DDM drift rate **(B, right)** without any effect on the boundary parameter **(B, left)**. Since the neural activity devoted to evidence accumulation (area under the curve) should increase **(C, left)**, we would expect higher BOLD signal in this case **(C, right)**. *Scenario 2: left SFS is causally involved in setting the choice criterion.* Theta-burst induced inhibition of left SFS should lead to a lower choice criterion **(D)**, expressed as lower choice accuracy **(D, 2nd row, left)**, faster RTs **(D, 2nd row, right)**, and a reduced DDM decision boundary parameter **(E, left)** without any effect on the DDM drift-rate **(E, right)**. At the neural level, we should observe reduced BOLD activity due to the lower amount of evidence

processed by the neurons **(F, right)**, and reflected by the smaller area under the evidence-accumulation curve when it reaches the lower boundary **(F, left)**.

RESULTS

The Experiment. We recorded functional magnetic resonance imaging (fMRI) data from hungry, healthy participants ($n = 20$) performing perceptual- and value-based choice-tasks in alternation (**Methods** and **Fig. 2b**). For perceptual decisions, participants chose the larger food item, while for value-based decisions, participants chose which food item they would prefer to receive and consume by the end of the experiment. The stimuli and motor responses were identical for both tasks. Choice pairings were predetermined based on participant's individual subjective perceptual- and value-based ratings of the food items, obtained just prior to the scanning session. Perceptual evidence was defined as the size difference (SD) between the food items, whereas value evidence was defined as the difference in value ratings (VD) between the choice alternatives (see **Methods** and **Fig. 2b**). A choice was classified as correct when it was consistent with the previously acquired ratings regarding size and preference respectively, i.e., when the larger-rated item was chosen for perceptual decisions or the higher-valued item was chosen for value-based decisions (Polania et al., 2014, 2015).

Before scrutinizing the role of the left SFS for either type of decision making, we first confirmed the validity of our choice paradigm behaviourally and neurally. Behavioral regressions confirmed that our experimental design allowed for a clear computational separation of both choice types: During perceptual decisions, participants relied exclusively on perceptual evidence, as reflected in both increased choice accuracy (main effect SD, $\beta = 0.560, p < 0.001$ and VD, $\beta = 0.023, p = 0.178$; **Fig. 2c**) and faster reaction times (RTs) with larger perceptual evidence but not value-based evidence (main effect SD,

$\beta = -0.057, p < 0.001$ and VD, $\beta = -0.002, p = 0.281$; **Fig. 2c**). Conversely, participants relied only on value evidence during value-based decisions, as evident from both choice accuracy (main effect VD, $\beta = 0.245, p < 0.001$ and SD, $\beta = -0.254, p = 0.123$; **Fig. 2c**) and RTs (main effect VD, $\beta = -0.016, p = 0.011$ and SD, $\phi = -0.003, p = 0.419$; **Fig. 2c**) irrespective of the items' size difference. Thus, our results replicate previous findings obtained with a similar paradigm (Polania et al., 2014, 2015; Grueschow et al., 2015) that participants can use exclusively task-relevant evidence to make choices, and they confirm the suitability of our paradigm for directly comparing perceptual and value-based decisions with matched stimuli and motor responses.

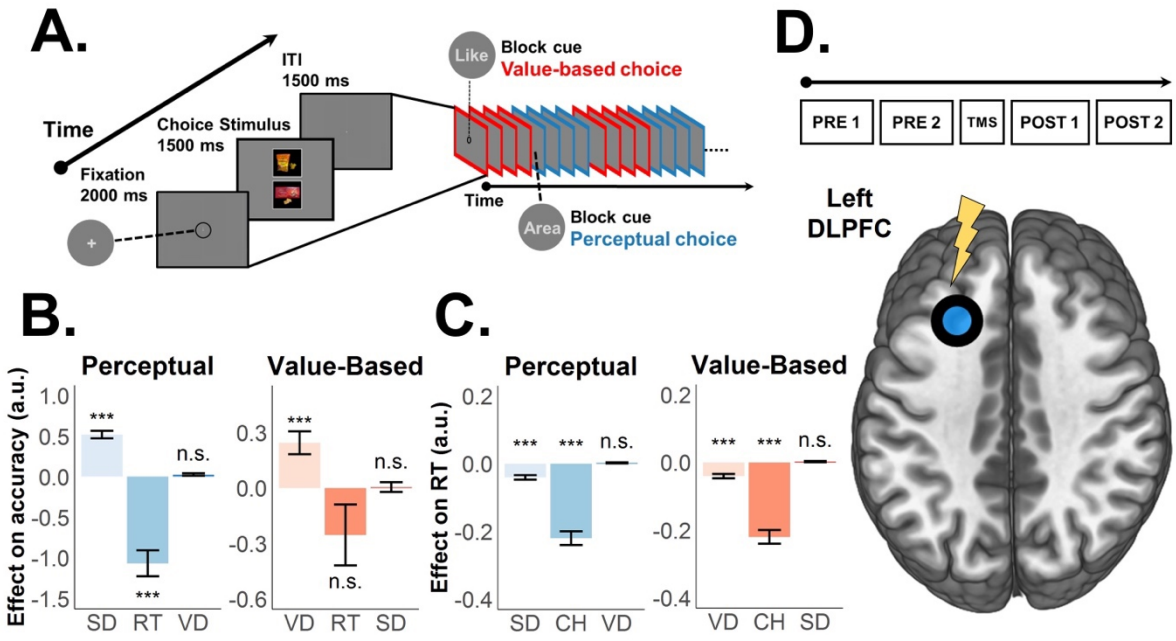

FIGURE 2. Behavioral food choice paradigm, theta-burst stimulation protocol, and behavioral regressions. (a) *Example of decision stage.* Participants were cued in advance about the type of decision required. Perceptual decisions required participants to choose the food item with the largest size while value-based decisions required participants to choose the food item they preferred to consume at the end of the experiment. Participants alternated between blocks of perceptual (blue) or value-based (red) choice trials (7-9 trials per task-block). **(b)** Regression results show that the larger the evidence strength, the more likely decision makers will respond accurately. Choice accuracy is only related to the evidence that

is currently task-relevant (size difference SD for perceptual or value difference VD for value-based choice), not to the task-irrelevant evidence (RT is reaction time of current choice) . **(c)** Similarly, we show that RTs are negatively associated only with the task-relevant evidence (and lower for perceptual choices overall, captured by regressor CH (1 = perceptual, 0 = value-based)). Consistent with previous findings, the results in **(b)** and **(c)** confirm that our paradigm can distinguish and compare evidence processing for matched perceptual- and value-based decisions. Error bars in **(b)** and **(c)** represent the 95% confidence interval range of the estimated effect sizes. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$. **(d)** *Theta-burst stimulation protocol.* After the fourth pre-TMS run, participants received continuous theta-burst stimulation (cTBS) over the left SFS region of interest (ROI) (area encircled and colored blue). cTBS consisted of 200 trains of 600 pulses of 5 Hz frequency for 50 s.

In line with the behavioural results that participants depended on different evidence for the two types of choices, initial fMRI analysis revealed that neural activations strongly differed between choice types (despite the fact that participants saw the same images and gave the same motor responses). In line with previous findings (Grueschow et al., 2015), we found that while visual and motor areas were jointly activated for both types of choices ($p < 0.05$, FWE-corrected with cluster forming thresholds at $T(19) > 2.9$; **Supplementary Fig 1a.** and **Supplementary Table 1**), perceptual decisions led to stronger recruitment of the posterior parietal cortex whereas value-based decisions led to stronger activations of the medial prefrontal cortex and posterior cingulate cortex (**Supplementary Fig 1b.** and **Supplementary Table 2**). These choice-type-specific brain activations, in response to identical visual input and motor output, ascertain that participants recruit task-specific brain regions depending on the choice domain.

Behavior: Theta-burst stimulation reduces choice accuracy only for perceptual decisions

Our experiment was divided into pre- and post-stimulation blocks. After participants had performed four pre-stimulation session-blocks inside the scanner, they received continuous

theta burst stimulation (cTBS) (Huang et al., 2005; Di Lazzaro et al., 2005, 2008) over the left SFS (MNI coordinates, $x = -24, y = 24, z = 36$; Heekeren et al., 2004; Philiastides et al., 2011; Grueschow et al., 2018). Following this intervention, participants completed four post-stimulation fMRI blocks. By comparing the effects of stimulation on both types of behavior and brain activity between post- and pre-stimulation blocks, we identify the role of SFS for either type of decision making. In particular, we examined whether the SFS is indeed selectively involved in perceptual decisions as previously suggested (Heekeren et al., 2004, 2006; Rahnev et al., 2016; Philiastides et al., 2011).

Our results support the hypothesis that the SFS has a specific role for perceptual decision making, on several experimental levels. Behaviourally, we found that SFS-cTBS led to a significant decrease from pre- to post-cTBS blocks in accuracy for perceptual choices (main stimulation effect, $\beta = -0.465 \pm 0.342$, $p = 0.008$; **Fig. 3a** and **Supplementary Fig. 2a**), while value-based choice consistency remained unaffected by SFS stimulation ($\beta = -0.042 \pm 0.205, p = 0.691$; **Fig. 3a** and **Supplementary Fig. 2a**). These differences were significant in direct comparison (stimulation × task interaction, $\beta = -0.094 \pm 0.087, p = 0.034$; **Fig. 3a; Supplementary Fig. 2c** and **Supplementary Table 2**). Interestingly, SFS-cTBS had comparable effects on reaction times in both tasks: Faster RTs were observed after SFS-cTBS for both perceptual (main stimulation effect, $\beta = -0.116 \pm 0.067, p = 0.003$; **Fig. 3b** and **Supplementary Fig. 2b**) and value-based choices (main stimulation effect, $\beta = -0.125 \pm 0.063, p = 0.001$; **Fig. 3b** and **Supplementary Fig. 2b**), with no significant difference between these two effects (stimulation × task interaction, $\beta = 0.009 \pm 0.069, p = 0.795$; **Fig. 3b; Supplementary Fig. 2c** and **Supplementary Table 2**)**.** These common changes in RTs from the first to the second half of the experiment may not reflect TMS-related changes in SFS function but rather general training effects common to both tasks (Hyman, 1953; Mowbray and Rhoades, 1959; Mawase

et al., 2018). We examined this possibility in more detail with computational modelling in the next section.

Modelling: SFS-TMS reduces decision boundary only for perceptual decisions

To examine in detail which specific latent decision process was affected by the SFS-cTBS, we fit the hierarchical drift diffusion model (HDDM) simultaneously to the accuracy and RT data of our participants. This canonical model of choices allowed us to identify and disentangle the effect of stimulation on various latent variables representing distinct components of the choice mechanism (Ratcliff, 1978; Ratcliff and Smith, 2004; Ratcliff and McKoon, 2008; Polania et al., 2015; **Supplementary Fig. 3** and see **Methods**). A specific focus of this analysis was on whether SFS-cTBS would change the way participants set the decision threshold (*boundary parameter;* Rahnev et al., 2016; Bogacz et al., 2010; Domenech and Dreher, 2010; Herz et al., 2016) or the efficiency with which choice-relevant evidence is accumulated (*drift-rate*, Philiastides et al., 2011; Basten et al., 2010) (see **Methods** for more details and **Figure 1b,e**).

Our results are consistent with a causal role for the left SFS in setting the decision threshold of perceptual choices: For perceptual decisions, our computational analysis reveals a significant post-cTBS decrease in boundary (see **Methods**; $p_{MCMC} = 0.003$; **Fig. 3c** and **Supplementary Fig. 5a**) but no such effect for any of the other parameters ($p_{MCMC} = 0.822$ for drift rate; **Fig 3d** and **Supplementary Fig. 5b,c**). For value-based decisions, by contrast, no effect of cTBS was observed for either of the two decision-relevant parameters ($p_{MCMC} = 0.115$ for boundary and $p_{MCMC} = 0.758$ for drift rate; **Fig. 3c,d** and **Supplementary Fig. 5a,b,c**), supporting the specificity of the SFS involvement in perceptual decisions. This conclusion was further corroborated by direct comparison of these effects, which showed that SFS-cTBS had a significantly stronger impact on the boundary parameter for perceptual compared to value-based decisions (stimulation × task interaction for the decision

threshold, $p_{MCMC} = 0.045$; **Supplementary Fig. 5a**; there were no such differences for drift-rate; $p_{MCMC} = 0.685$; **Supplementary Fig. 5b**).

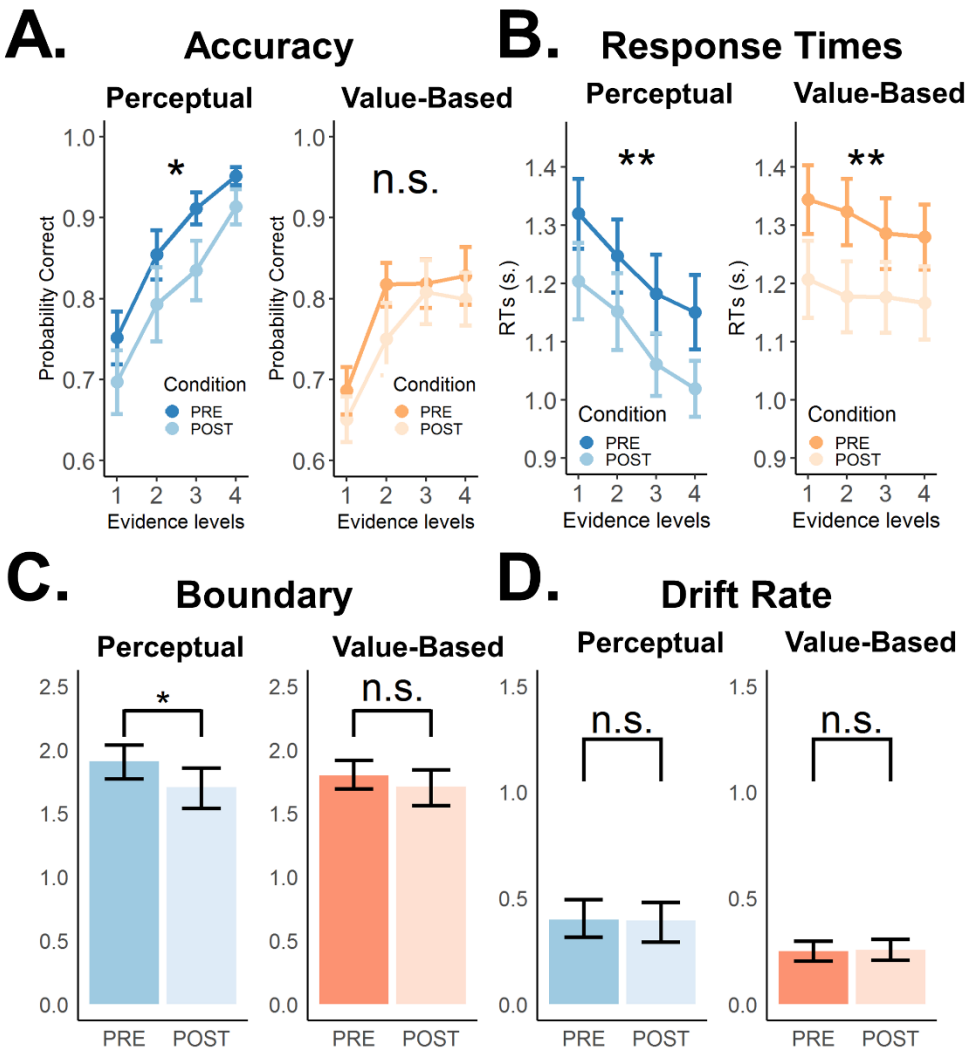

FIGURE 3 Theta-burst stimulation over the left SFS affects choice behavior and selectively lowers the decision boundary for perceptual but not value-based choices. **(a)** Accuracies and **(b)** response times (RTs) for perceptual (blue) and value-based (orange) decisions for different evidence levels during pre-cTBS (dark) and post-cTBS (light) stimulation periods. Error bars in **(a)** and **(b)** represent s.e.m. Consistent with previous findings, stronger evidence leads to more accurate choices and faster RTs in both types of decisions. Importantly, theta-burst stimulation significantly lowered choice accuracy selectively for perceptual, not value-based decisions (negative main stimulation effect for perceptual decisions and negative stimulation × task interaction; **Supplementary Fig. 2c**

and see also **Supplementary Fig. 2a** for changes in choice accuracy across runs). Additionally, theta-burst stimulation also significantly lowered RTs in both choice types (negative main stimulation effect; **Supplementary Fig. 2c** and see also **Supplementary Fig. 2b** for changes in RTs across runs). **(c)** Theta-burst stimulation selectively decreased the decision boundary in perceptual decisions only (difference between estimated posterior population distributions; see **Methods** and **Supplementary Fig. 5a** for a detailed post-hoc analysis). All the other parameters, particularly **(d)** the drift rate (see also **Supplementary Fig. 5b** for post-hoc analysis) remain unaffected by stimulation. Error bars in **(c)** and **(d)** represent the 95% confidence interval range of the posterior estimates of the DDM parameters. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

Examining other DDM parameters showed that the faster RTs observed for value-based decisions after the stimulation did not reflect evidence-dependent choice processes, but rather a change in non-decision-related sensorimotor processes (nDT) (see **Methods**; **Supplementary Fig. 3**), a DDM parameter that indexes constant latencies associated with sensory and motor preparation processes that are invariant across trials with different choice evidence (Verdonck and Tuerlinckx, 2016; Starns and Ma, 2017). This parameter was marginally decreased after stimulation for value-based ($p_{MCMC} = 0.062$) but not perceptual decisions ($p_{MCMC} = 0.707$) (**Supplementary Fig. 4b** and **5c**), with a significant difference between these effects ($p_{MCMC} = 0.041$; **Supplementary Fig. 5c**). A decrease in nDT is unrelated to changes in the evidence accumulation process (Feltgen and Daunizeau, 2020; White et al., 2018) and may therefore reflect learning processes (Hyman, 1953; Mawase et al., 2018).

fMRI: SFS activation changes for perceptual choices in line with model predictions

To investigate whether our behavioural and computational results directly relate to task-specific disruption of neural activity in left SFS, we investigated BOLD response changes in this brain area after stimulation. We exploited the fact that our fitted DDM and its latent

parameters make clear predictions about how BOLD responses in this area should change if the stimulation affects the neural computations involved in setting the boundary for the necessary amount of evidence accumulation. Importantly, these predictions translate to clear parametric regressors that we can use for trialwise analysis of fMRI data (Basten et al., 2010; Domenech et al., 2018; Liu and Pleskac, 2011). More specifically, we expected that the BOLD signal level is proportional to the DDM's accumulated evidence (*aE*), defined as the area below the modelled evidence accumulation curve up until the accumulator reaches the decision boundary (Liu and Pleskac, 2011; Domenech et al., 2018; Basten et al., 2010). Using subject-wise DDM latent parameters, the average area below the decision boundary for each evidence level can be computed as a function of each *participant's decision boundary divided by the mean drift rate* (see **Fig. 1c and 1f** and **Methods** for more details). Using the more detailed trialwise measures, however, the same area can be computed as a function of each *trial's RTs divided by the evidence level*, since according to the DDM, the duration of response times is directly proportional to the decision boundary, and the evidence level is directly proportional to the slope of the drift rate (Ratcliff and Rouder, 1998; Ratcliff and McKoon, 2008; see **Methods** for more details). Exploiting these two known facts from the DDM thus allows us to extend our test of the stimulation effect from individual-specific latent parameters to trialwise regressors and behavioral measures. Higher SFS BOLD signals are associated with higher *aE* and vice versa (Basten et al., 2010; Liu and Pleskac, 2011; Filimon et al, 2013; Tosoni et al., 2009), implying that a TMS intervention lowering the decision boundary should lower *aE* and therefore BOLD signals. Crucially, these latent changes predicted by the DDM should also be reflected in the subject-level simulations of accumulated evidence constructed from the DDM parameters.

Thus, we first tested whether our neural hypotheses would already be evident in the simulated trial-wise *aE* regressors. We used individual parameters identified by fitting our computational framework to simulate expected neural activity on a trial-wise basis across

participants. To this end, we derived the predicted *aE* from the model parameters for each participant. A comparison across cTBS and task conditions confirmed the predicted cTBS-related decrease in accumulated perceptual evidence for perceptual decisions ($p_{mcmc} = 0.003$; **Fig. 4a** and **Supplementary Fig. 6b**), the corresponding null effect for value-based decisions ($p_{mcmc} = 0.100$; **Fig. 4b** and **Supplementary Fig. 6b**), and a significant difference for this effect between both choice types (one-sided $p_{mcmc} = 0.048$; **Supplementary Fig. 6b**).

In the next step, we used the *trial-by-trial accumulated evidence* as a regressor in the statistical analysis of the BOLD signals, allowing us to test whether the left SFS shows the predicted changes in neural response to varying levels of perceptual evidence. First, we tested whether our predictor of neural *accumulated evidence* was represented in BOLD signals of similar task-specific areas as reported previously for perceptual choices in SFS (Heekeren et al, 2004, 2006) and for value-based choices in vmPFC (De Martino et al, 2013; Grueschow et al, 2015). This was confirmed by the data: During perceptual choices, trialwise *aE* correlated with BOLD activity in the left SFS (peak at $x = -21, y = 26, z = 37$; SVC < 0.05; **Supplementary Fig. 7b** and **Supplementary Table 3**) whereas during value-based choices, it related to BOLD activity in the ventromedial prefrontal cortex (vmPFC) (peak at $x = 3, y = 38, z = -17$; $SVC < 0.05$; **Supplementary Fig. 7e**) and the nucleus accumbens (peak at $x = 9, y = 11, z = -11$; $p < 0.05$, FWE-corrected with cluster-forming thresholds at $T(19) > 2.9$; **Supplementary Fig. 7e**). For both types of choices, domain-general representations of *aE* were also evident (see **Supplementary Fig. 7** and **Supplementary Table 3**).

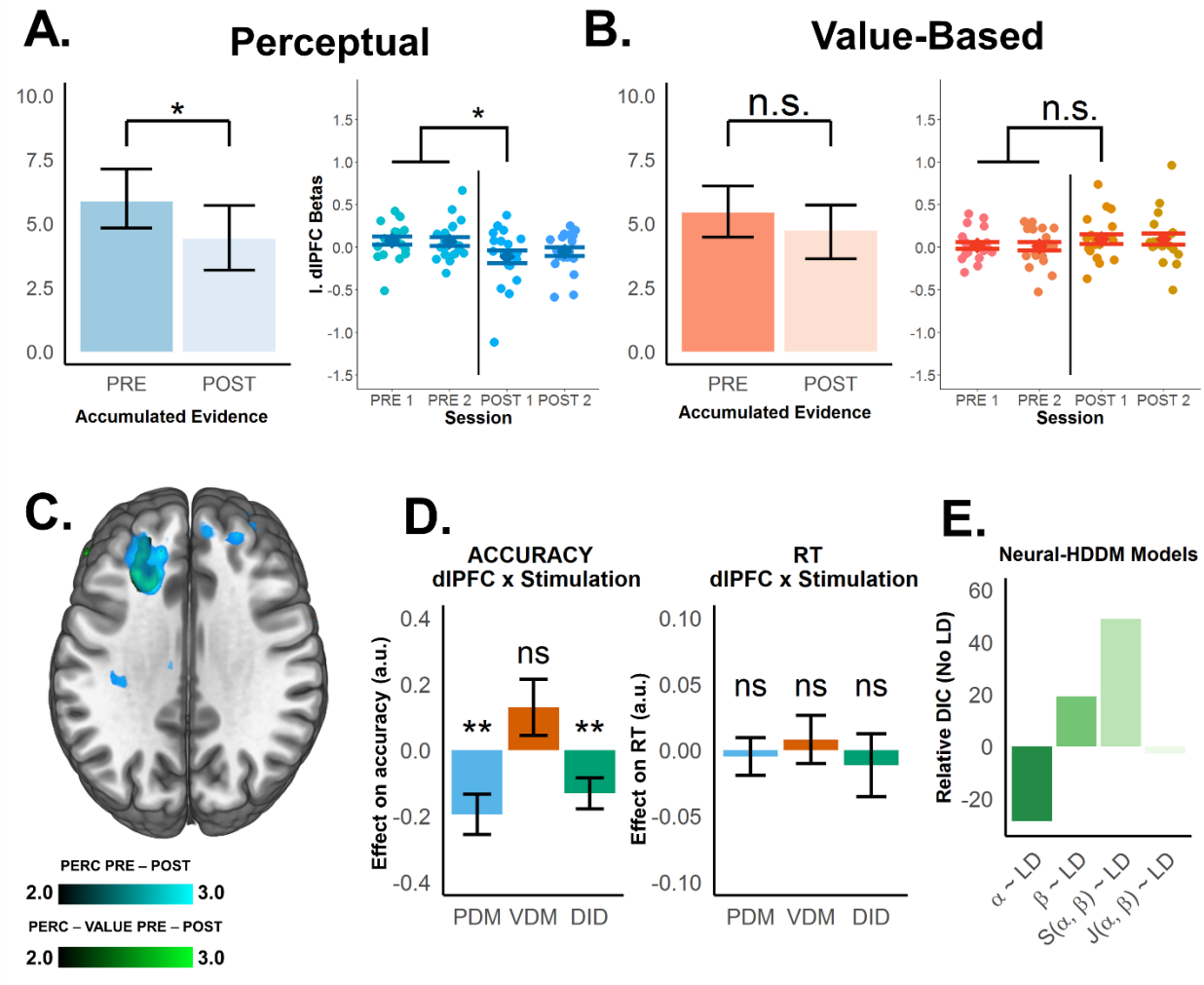

FIGURE 4. Neural representation of accumulated evidence in the left SFS is disrupted after theta-burst stimulation, and is linked with behavior and neural computation. (a) Left panel: Accumulated evidence (*AE*) simulation derived from the fitted DDM (left panel). Previous studies have illustrated how the accumulation-to-bound process convolved with the hemodynamic response function (HRF) results in BOLD signals; hence, the simulated *AE* provides a suitable prediction of BOLD responses in brain regions involved in evidence accumulation. Theta-burst stimulation selectively decreased *AE* for **(a)** perceptual (blue), not **(b)** value-based (orange) decisions (see **Supplementary Fig. 6b** for post-hoc analysis). We constructed a trialwise measure of accumulated evidence using RTs and evidence strength for our parametric modulator (see **Methods**). Individual ROIs extracted from the left SFS representing accumulated evidence across runs (right panels; see **Methods**) show that consistent with the DDM prediction, theta-burst stimulation selectively decreased BOLD response representing *AE* in left SFS during perceptual, not value-based decisions. Error bars

in the left panels of **(a)** and **(b)** represent the 95% confidence interval range of the posterior estimates of the DDM parameters, while error bars in their respective right panels represent s.e.m. **(c)** Comparing pre- and post-cTBS contrasts of BOLD signals related to accumulated evidence, during perceptual decisions, show signal changes in left SFS (green) after theta-burst stimulation. Further contrasts comparing pre-post difference across both choice types (blue) confirm the selectivity of TMS effects for perceptual decisions. **(d)** To test the link between neural and behavioural effects of TMS, regression results show that after stimulation, BOLD changes in left SFS are associated with lower choice accuracy (left panel) for perceptual (PDM, blue) (negative left SFS × stimulation interaction) but not value-based choices (VDM, red), with significant differences between the effects on both choice types (difference-in-difference, DID, green, negative left SFS × stimulation × task interaction). On the other hand, cTBA-induced changes in left SFS activity are unrelated to changes in RT (right panel). Error bars in **(d)** represent the 95% confidence interval range of the estimated effect sizes. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$. **(e)** To test the link between neural activity and DDM computations, we included trialwise beta estimates of left-SFS BOLD signals as inputs to the DDM. Alternative models tested whether trialwise left-SFS (LD) activity modulates the decision boundary (α) (Model 1), the drift rate (β), or a combination of both (Models 3 and 4, see **Methods** and **Supplementary Fig. 8** for more details). Model comparisons using the deviance information criterion (DIC, smaller values mean better fits) showed that Model 1 fits the data best, confirming that the left SFS is involved in selectively changing the decision boundary for perceptual decisions.

We then tested whether the cTBS specifically reduced the neural representation of accumulated perceptual evidence in the left SFS for perceptual decisions, as predicted by the behavioral and modelling results. In line with these predictions, comparison of the post – pre trial-aE regressor showed a lower BOLD response in left SFS to the trialwise perceptual evidence during perceptual decisions ($SVC < 0.05$; **Fig. 4c, green patch**). This effect was significantly stronger than the corresponding effect on evidence representations in this area during value-based decisions ($SVC < 0.05$; **Fig. 4c, blue patch**). No effect was found for value-based decisions alone. Convergent evidence for the specificity of this effect was provided by an alternative hypothesis-guided region-of-interest (ROI) analysis of the

regression weights extracted from an a-priori ROI-Mask of the SFS (see **Methods**). This showed lower post-stimulation beta values for the trial-aE regressor during perceptual (main stimulation effect, $\beta = -0.153 \pm 0.054, p = 0.004$; **Fig. 4a**) but not value-based choices (main stimulation effect, $\beta = 0.078 \pm 0.053, p = 0.140$; **Fig. 4b**) and a significant difference in these effects (stimulation × task interaction, $\beta = -0.232 \pm 0.075, p = 0.002$; **Fig, 4a,b**). Thus, the fMRI results show that cTBS of the left SFS indeed affects neural processing in this brain structure selectively during perceptual choices, in a way that is consistent with a lowering of the boundary and less accumulated evidence as predicted by the fitted DDM model. This remarkable convergence between the behavioural, modelling, and fMRI results suggests that the left SFS is indeed causally involved in setting decision criteria for choices based on perceptual evidence, but not based on subjective values.

fMRI: Perceptual-choice accuracy and boundary setting reflect trial-by-trial changes in SFS activity

If perceptual-decision performance depends specifically on activity in the left SFS, then trial-wise choice accuracy should relate to trial-wise BOLD activity in the SFS during perceptual decisions, over and above the mean effects of evidence level. To test this, we regressed choice accuracy on trial-by-trial BOLD activity extracted from the left SFS ROI, choice type, and TMS, while controlling for the evidence provided by the stimulus pairs on each trial (see **Methods** for details). In line with our prediction, we observed that the relation between perceptual-choice accuracy and trial-by-trial SFS activity was significantly decreased by TMS (SFS × stimulation interaction, $\beta = -0.196 \pm 0.128, p = 0.003$; **Fig 4d**), independently of the corresponding effects for choice evidence (SD main effect, $\beta = 0.524 \pm 0.082$, $p < 0.001$, VD main effect, $\beta = 0.197 \pm 0.012, p = 0.001$, SFS × SD interaction, $\beta = -0.041 \pm 0.046$, $p = 0.365$, SFS × VD interaction, $\beta = 0.055 \pm 0.041, p = 0.183$). This effect was clearly specific for perceptual decisions, since no such effects were observed for

value-based choices (SFS × stimulation interaction $\beta = 0.099 \pm 0.242, p = 0.422$; SFS × stimulation × task interaction, $\beta = -0.072 \pm 0.051, p = 0.005$; **Fig 4d**) and for RTs during both types of choices (SFS × stimulation interaction, perceptual: $\beta = -0.031 \pm 0.053, p = 0.367$; **Fig 4d;** accuracy: SFS × stimulation interaction, $\beta = -0.012 \pm 0.050, p = 0.650$; **Fig 4d**).

We further investigated whether the relation between trialwise SFS activity and choice outcome indeed reflected an SFS role for perceptual boundary setting, as suggested by the DDM results presented above. To confirm this neurally, we set up several DDMs with trialwise SFS activity as an additional modulator for DDM parameters (on top of choice evidence; see methods and Herz et al., 2016, 2017; Turner et al., 2015). More specifically, we tested several DDMs in which trialwise SFS activity either modulated the decision threshold only (**Model 1; Supplementary Fig. 8a**), the drift rate only (**Model 2; Supplementary Fig. 8b**), or both parameters separately (**Model 3; Supplementary Fig. 8c**) or jointly (**Model 4; Supplementary Fig. 8d**). We compared these neural HDDMs to our baseline HDDM without neural inputs (see **Methods** for more details and **Supplementary Fig. 3**), allowing us to test across all conditions and choice types whether model evidence was enhanced when adding a potential trial-by-trial influences of SFS activity to the experimental inputs. Thus, the reported model evidence criterion (DIC) provides an additional formal test of whether the cTBS-influenced SFS activity relates selectively to the decrease of the decision boundary for perceptual choices only. Consistent with this prediction, **Model 1** where SFS activity modulated the decision threshold only, outperformed all other models and model evidence showed improvements versus the baseline model (relative $DIC = -28.65$; **Fig. 4b**). These results provide direct evidence that neural computations in the left SFS support criterion setting for perceptual evidence accumulation.

fMRI: TMS affects SFS functional connectivity during perceptual choices.

Our results so far indicate that cTBS to the left SFS disrupts selectively a neural process related to setting the criterion for perceptual evidence accumulation. However, it is conceivable that SFS-cTBS may also change functional communication of the SFS with other brain areas involved in initial processing of the perceptual information necessary to make a choice. We investigated this possibility by testing whether cTBS affected functional coupling of the SFS. A psychophysiological interaction (PPI) analysis seeded in left SFS and modulated by aE indeed revealed stronger coupling with occipital cortex (OCC) after cTBS (peak at $x = -28, y = -85, z = -2$; $p < 0.05$, FWE-cluster-forming thresholds at $T(19) > 2.9$; **Fig. 5a**). Interestingly, the activity peak in visual cortex showing evidence-dependent coupling with SFS, overlaps with the spatiotopic neural representation of the stimulus items in the visual field during decision making. We identified this overlap using a conjunction analysis of the PPI result and a contrast regressing BOLD signal on trial-by-trial stimulus onsets of both choice types (at familywise-error-corrected thresholds). Moreover, we used the latter contrast to define fully independent regions-of-interest (ROIs) in occipital cortex processing the visual stimuli independent of task type and performed an ROI analysis on the individual SFS-OCC-PPI betas extracted for each participant. This confirmed that evidence-related functional coupling is increased by stimulation during perceptual (main stimulation effect, $\beta = 0.330 \pm 0.284, p = 0.022$) but not value-based choices (main stimulation effect, $\beta = -0.186 \pm 0.247, p = 0.139$; **Fig. 5b;** stimulation × task interaction, $\beta = 0.517 \pm 0.44, p = 0.021$; **Fig. 5a**). Thus, our results indicate that cTBS to the left SFS leads to stronger functional coupling with occipital areas involved in processing the visual stimuli, perhaps consistent with increased downstream demand on visual-related resources when upstream evidence accumulation regions are impaired.

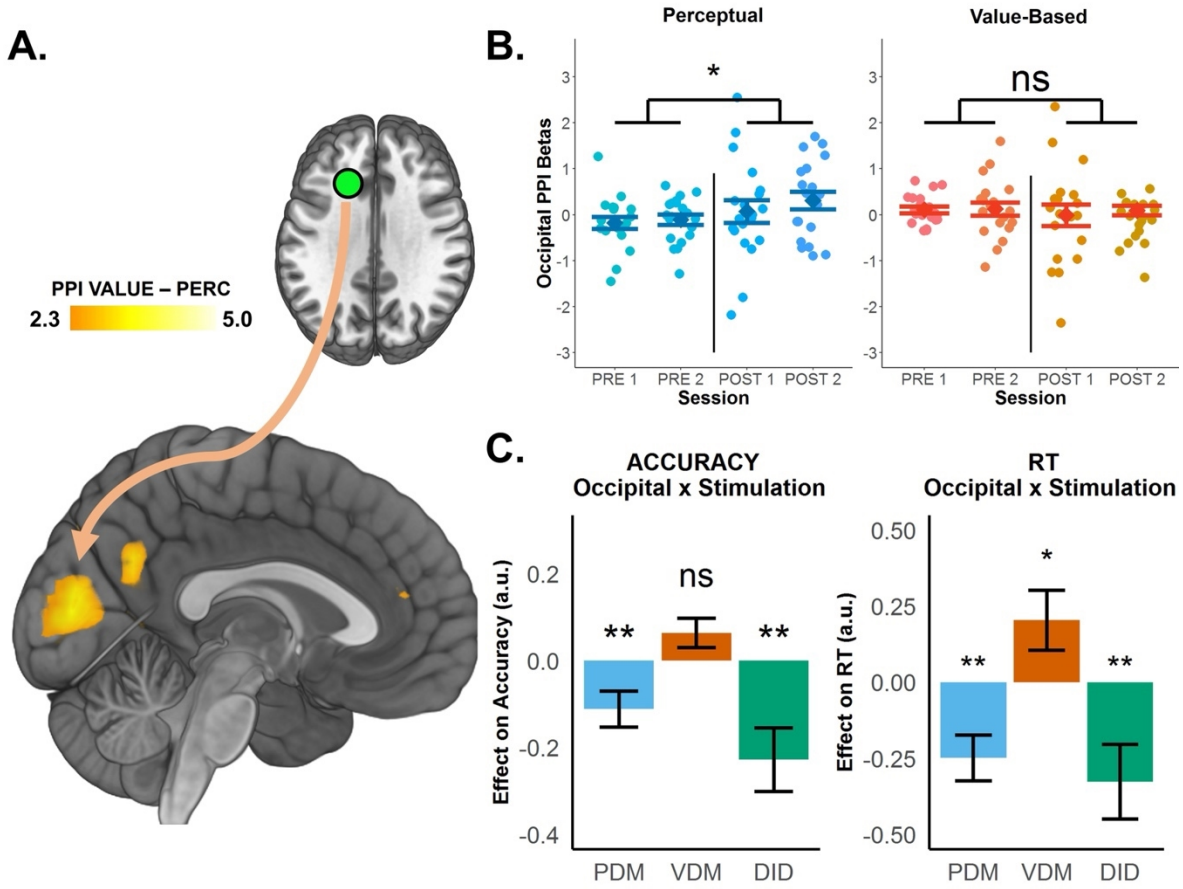

FIGURE 5. SFS-TMS-related changes in behaviour and neural computations are accompanied by increased functional coupling between the left SFS and occipital cortex. (a) Psychophysiological interaction (PPI) analysis reveals an area in occipital cortex showing increased functional coupling with the left SFS during perceptual choices. **(b)** ROI analysis of individual PPI betas shows that *aE*-related functional coupling between the left SFS and OCC is selectively increased post stimulation during perceptual (left panel) but not value-based decisions (right panel). Error bars in **(b)** represent s.e.m. **(c)** Regression results testing the link between cTBS effects on left SFS-OCC functional coupling and behaviour. Increased SFS-OCC coupling is associated with lower choice accuracy (left panel) specifically for perceptual (PDM, blue, negative OCC × stimulation interaction) but not value-based choices (VDM, red). In addition, increased functional coupling is also associated with faster RTs (right panel) for perceptual (blue, negative OCC × stimulation interaction) and slower RTs for value-based choice (red, positive OCC × stimulation interaction). Error bars in **(c)**

represent the 95% confidence interval range of the estimated effect sizes. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

We further explored whether this TMS-induced increase in functional coupling between the left SFS and OCC is related to changes in behavior and specific neural computations during perceptual decisions. To test this, we related these effects to individual measures of choice behavior and latent DDM parameters for each participant. This revealed that stimulation-induced increases in SFS-OCC coupling were associated with lower accuracy (OCC $\times$ stimulation $\times$ task interaction, $\beta = -0.225 \pm 0.142, p = 0.002$; **Fig. 5c**) and shorter RTs (OCC $\times$ stimulation $\times$ task interaction, $\beta = -0.325 \pm 0.238, p = 0.007$; **Fig. 5c**) for perceptual, but not value-based decisions. Taken together, these results thus show that the causal behavioral and computational changes during perceptual decisions due to left SFS TMS relate not just to local neural changes in the SFS, but also to the way this brain structure communicates with visual cortex.

DISCUSSION

We investigated whether the left SFS serves a domain-specific or domain-general role in decision making, and which precise functional contribution it makes to the choice process. Previous work had implicated the left SFS in both perceptual and value-based decision making (Basten et al., 2010; Hare et al., 2011; Heekeren et al., 2004), but direct comparison between these choice domains were largely missing from the literature. Moreover, the precise role of the SFS has been debated, with contradictory results that either reported a role in setting the decision threshold (Rahnev et al., 2016) or tracking the evidence strength of incoming sensory information (Philiastides et al., 2011). Our findings resolve this uncertainty regarding the nature and specificity of left-SFS involvement in decision making,

by showing that it is causally involved in setting the criterion during evidence accumulation only for perceptual but not value-based choices.

Implications for theories of PFC organization. Our results directly inform long-standing debates about functional specialization versus integration in the prefrontal cortex (PFC). While it is widely agreed that the PFC is not a homogenous region (Goldman-Rakic, 1984, 1994; Owen, 1997; Pandya and Barnes, 1987), considerable debate has surrounded the principles by which it is functionally organized (Badre and D'Esposito, 2007; Owen, 1997; Goldman-Rakic, 1995; Nee and D'Esposito, 2016; Koechlin, 2003; Petrides, 2005). *Process-specific* and *domain-general models* posit that different PFC regions may contribute to specific aspects of information processing, in a manner that that can be flexibly applied to all types of information, be it from different sensory modalities or in different cognitive formats (Owen et al., 1996; Petrides, 1995, 2005). On the other hand, *domain-specific models* suggest that PFC regions are fractionated and functionally organized to process specific types of information, as determined based on their anatomical inputs (Goldman-Rakic, 1984, 1995; Goldman-Rakic and Leung, 2002; Levy and Goldman-Rakic, 2000; Snow, 2016). Our results are more consistent with the latter domain-specific account, since they revealed that causally interfering with the left SFS selectively affects perceptual decisions, leaving performance during value-based decisions unaffected.

Moreover, while the PFC is implicated in various specific cognitive functions (Duncan and Owen, 2000; Meng et al., 2016; Raos and Savaki, 2016), prevailing perspectives have also characterized the PFC as an anterior-to-posterior hierarchy organized for the purpose of general cognitive control and executive function (Koechlin, 2003; Koechlin and Jubault, 2006; Koechlin et al., 1999; Nee and D'Esposito, 2016). This view suggests that the main role of the PFC is largely in the domain of higher-order cognitive and abstract operations that transcend specific functional domains (Domenech and Koechlin, 2015; Koechlin and Summerfield,

2007; Owen, 1997). However, an alternative account suggests that the left SFS plays a specialized role in the selective accumulation process for low-level perceptual evidence (Heekeren et al., 2004, 2006, 2008; Philiastides et al., 2011; Rahnev et al., 2016). Our results are more consistent with the latter view, as we provide not only causal evidence, but also demonstrate specificity for processing perceptual evidence in the SFS.

Do value-based decisions also rely on distinct PFC areas? Previous work has suggested that during value-based decisions, the DLPFC interacts with the vmPFC in modulating the value signal to facilitate self-control (Hare et al., 2009; Rudorf and Hare, 2014; Maier et al., 2015). Our results do not contrast with these previous findings: The DLPFC sub-regions reported by these previous self-control studies are found around the inferior and middle frontal gyri (Hare et al., 2009; Rudorf and Hare, 2014; Maier et al., 2015) and thus in PFC areas that are anatomically distinct (and in fact distant) from the superior frontal sulcus. This further underlines that our findings are most consistent with a functional organization of PFC as a collection of fractionated sub-regions, where each region processes different types of information (Goldman-Rakic, 1995; Goldman-Rakic and Leung, 2002; Levy and Goldman-Rakic, 2000). Future studies should consider addressing this issue by causally targeting these specific regions in frontal gyri and compare whether it affects only the processing of self-control during value-based decisions, or whether it may affect perceptual decisions as well.

Our study used a well-established two-alternative forced-choice design with matched responses and relatively simple decisions (Polanía et al., 2014, 2015; Grueschow, 2018). However, the value-based choice domain entails a large array of choice types with varying degrees of complexity. For instance, more complex types of value-based decisions entail decisions under risk (Stewart et al., 2016; Glickman et al., 2019), intertemporal choice (Peters and D'Esposito, 2020), and strategic and social decisions (Hutcherson et al., 2015; Bottemanne and Dreher, 2019), which may plausibly recruit the PFC due to working memory

demands (Barbey et al., 2013; Skagerlund et al., 2016), adjustment of decision time (Sokol-Hessner et al., 2012), or cost-benefit computations (Basten et al., 2010). In light of these many additional decision types based on preferences in the value-based choice domain, the functional specificity of SFS to perceptual decisions we claim here may have to be viewed with caution. Future studies should consider exploring the comparison between more complex types of value-based decisions with perceptual decisions, while taking great care in matching the degree of complexity between the two choice domains to avoid confounds induced by context or task difficulty.

Specificity of the SFS for perceptual decisions – only in humans? Our finding of a selective role of left SFS in perceptual evidence accumulation is particularly intriguing. The area appears to be uniquely developed in the human brain, with no close anatomical homologue in other species. In the animal literature, most prefrontal disruption studies in non-human primates have focused on the frontal eye fields (FEF) (Ding and Gold, 2012a; Hanks and Summerfield, 2017; Shadlen and Newsome, 1996) and in rodents on the frontal orienting fields (FOF) (Erlich et al., 2015; Hanks et al., 2015). While we and others observed disruption of the evidence accumulation process after interfering with SFS function in humans (Philiastides et al., 2011; Rahnev et al., 2016), disruption of the FOF in rodents has not affected behavior at all or in a qualitatively different manner (Brody and Hanks, 2016; Erlich et al., 2015; Hanks et al., 2015). However, the results of electrical stimulation of the FEF in monkeys (Ding and Gold, 2012a; Hanks and Summerfield, 2017) cannot necessarily be directly compared with TMS studies of human SFS, since FEF and SFS in humans are both structurally and functionally distinct (Murd et al., 2020; Rahnev et al., 2016). Thus, while it is tempting to speculate that the SFS perceptual evidence accumulation process identified here may be specific to humans, it is possible that researchers may have to further consider other

putative homologues across species that may truly correspond to the SFS area stimulated here (Brunton et al., 2013; Hanks and Summerfield, 2017).

Functional coupling between the left SFS and visual cortex. Our results suggest that the SFS role for domain-specific accumulation of perceptual evidence is not just a local phenomenon but extends to functional communication with visual areas. It is well-established that the prefrontal cortex is structurally connected with many other brain regions (Wycoco et al., 2013) and may flexibly interact functionally with different areas depending on choice demands. For example, for value-based decisions, functional coupling has been established between the left prefrontal cortex and orbitofrontal cortex (Hare et al., 2011; Sokol-Hessner et al., 2012), motor cortex (Filimon et al., 2013; Goldman and Nauta, 1977; Tosoni et al., 2008), hippocampus (Courtney et al., 1998; Meng et al., 2016; Rowe et al., 2000), and parietal cortex (Polanía et al., 2014). For perceptual decisions, numerous studies have shown that the frontal eye fields in humans and non-human primates are also functionally coupled with areas in visual cortex (Armstrong et al, 2009; Armstrong and Moore, 2014; Barbas and Mesulam, 1981; Goldman-Rakic, 1987; Cameron et al., 2015; Curtis and D'Esposito, 2006; Ruff et al., 2006). Our results are generally consistent with an occipito-frontal information exchange but extend it specifically to the SFS during perceptual evidence accumulation (Bullier et al., 1996; Jao Keehn et al., 2019). Inhibition of this area's functional contribution to evidence accumulation led to an increase in its functional coupling with areas in occipital cortex representing the stimuli visually upon which choices were based. The changes in functional coupling strength between the two cortical regions corresponded to observed behavioral and latent computational changes. This suggests that perceptual choices rely not only on local processing in SFS but on an integrated functional circuit, comprising both SFS and occipital cortex, at least for decisions based on visual stimuli as studied here. Future studies should test whether perceptual choices based on other sensory modalities

(e.g., touch, audition) lead to a flexible coupling of SFS with the specific sensory areas processing these stimuli.

We can speculate why the occipital cortex may have been recruited after inhibition of the left SFS via cTBS stimulation. For example, it is possible that cTBS-related impairments in the accumulation mechanism implemented by the SFS biases the system to rely on second-best suboptimal mechanisms for solving the tasks, such as template matching from working memory. Previous work has provided converging evidence that maintenance of visual information in working memory enhances coupling between sensory processing in visual cortex and information storage in lateral prefrontal cortex (Gazzaley et al., 2004, 2007; Postle et al., 2000; Serences et al., 2009). In fact, it has been suggested that the dorsolateral PFC is canonically organized in "memory receptive fields" (Postle, 2016) that may be more heavily taxed when direct accumulation mechanisms for sensory input are impaired, as in the case of cTBS manipulations. Of course, there are many other candidate mechanisms, such as attention or working memory, that may be more heavily taxed to compensate for the excitability manipulation of the SFS area specialized for processing the sensory evidence, as suggested by previous work on prefontal-occipital interactions during various attention and working memory tasks (Awh and Jonides, 2001; Awh et al., 2006; Gazzaley et al., 2007; Zanto et al., 2011). In any case, our study shows clearly that in the healthy, undisrupted human brain, left SFS plays a key role in transforming perceptual evidence into choices.

A specific role of left SFS for criterion setting during perceptual decisions. Previous work has shown that the SFS activity correlates with the evidence strength in the accumulation process, as reflected by the drift rate (Basten et al., 2010; Heekeren et al., 2004, 2006). In support of this notion, cortical activity disruption with transcranial magnitude stimulation (rTMS) resulted in lower choice accuracy and slower RTs (Philiastides et al., 2011). However, our findings support another view, namely that left SFS is causally involved

in modulating the decision threshold, with clearly consistent results across behavioral, computational, neural, and neural-behavioral levels. Thus, our findings are more in line with findings by Rahnev et al. (2016), who also suggested that SFS disruption leads to lower threshold setting. It is important to note that the results of all these studies are not entirely incompatible, since impairments in both boundary and drift account lead to lower choice consistency (Cavanagh et al., 2011; Domenech and Dreher, 2010; Georgiev et al., 2016; Green et al., 2012; Herz et al., 2016; Philiastides et al., 2011). The main difference in fact concerns reaction times: lower drift rate implies slower RTs while a lower boundary implies faster RTs. One may speculate whether this divergence in results indeed reflects fundamentally different functional contributions or mainly differences in task design, leading to different computational demands in different contexts. For instance, the study by (Philiastides et al., 2011) presented dynamic series of briefly presented sequential stimuli, varying the strength of each stimulus within noise to vary the evidence levels. By contrast, our study and that by Rahnev (2016) presented one static stimulus pair *simultaneously* during a two-alternative forced choice and varied evidence not by noise but by stimulus difference. These are fundamentally different task designs that may plausibly account for the differences in our findings: Dynamic stimulus streams demand quick, sequential choices on each of the noisy brief stimuli; longer reaction times may therefore reflect that the ability to filter out visual noise is reduced. In contrast, simultaneous static stimulus presentation requires the decision maker to discriminate the evidence contained in two clearly-presented choice alternatives, which may place much larger demands on a correct setting of the threshold for deciding whether one stimulus is larger than the other.

In line with these considerations, previous studies have consistently shown that simultaneous versus sequential decisions have different effects on behavior (Ahmad et al., 2017; Mogilner et al., 2013; Ricker and Cowan, 2014) and recruit different neural processing (Ballesta and Padoa-Schioppa, 2019; Ditz and Nieder, 2020; Zhang et al., 2013). Moreover,

the presentation of one or more competing stimuli would result in different degrees of adjustment in visual receptive fields via stimulus suppression (Beck and Kastner, 2007; Luck et al., 1997), which may potentially recruit a different evidence accumulation processes as opposed to when there is only one stimulus presented at a time. Future studies should consider addressing this issue directly, by comparing within-subjects effects of SFS TMS on behavior, computation, and neural processing for simultaneous versus sequential stimulus presentations.

Implications for decision neuroscience. Many human decision neuroscience studies have employed model-based approaches to identify BOLD signals that correspond to computational processes (Carandini, 2012; Forstmann and Wagenmakers, 2015; Forstmann et al., 2011; Marr, 2010; Palmeri et al., 2017; Wijeakumar et al., 2017). However, the links between neural and latent computational processing established by these studies is largely correlational (Logothetis, 2008; Poldrack, 2006; Ramsey et al., 2010), and there are many model alternatives that could possibly account for BOLD signals. Our study illustrates that causal manipulations induced by targeted functional inhibition of brain areas can provide decisive information and provide more direct support for neurocomputational mechanisms posited by cognitive models. Specifically, our study underlines that the DDM provides a plausible mechanistic account of the decision process (Herz et al., 2016, 2017; Turner et al., 2015), by showing that left SFS inhibition by cTBS affects the evidence representation posited by the model consistently across behavioral, computational, neural, neural-behavioral levels. Importantly, our results directly link changes in behavior to changes in both latent computations and neural processing, by demonstrating how raw trialwise neural signals from the left SFS can augment the DDM to explain behavior. This suggests that once brain stimulation studies have established (causal) correspondence between neural activity and latent variables in decision models, such models can be fruitfully extended by neural

measures to provide a more complete characterization and prediction of choice behavior and potentially its malfunctions.

Implications for computational psychiatry. Accumulating evidence to a sufficient decision threshold is crucial for maintaining accurate decisions, and deficiencies in areas involved in threshold setting will likely lead to maladaptive behavior with potentially lifelong consequences (Beck et al., 2009; Gatto and Aldinio, 2019; Green et al., 2012; Herz et al., 2016, 2017; Ide et al., 2017; Li et al., 2009; Specker et al., 1995). Manifestations of impulsive behavior (Heyes et al., 2012; Moeller et al., 2001) are largely apparent in clinical populations with aberrations in decision threshold setting (Herz et al., 2014,2016). However, most studies of these disorders have focused on impulsive behavior induced by reward or preferences (Barack and Platt, 2017; Glimcher et al., 2007; Mäntylä et al., 2012; Mcclure and Bickel, 2014; Mischel et al., 1972). It is important to note here that reward impulsivity is only one of the many domains of aberrant behavior in clinical populations. Perceptual impulsivity can also be important, since many of the behavioral and cognitive deficits are closely linked to impairments in perceptual function (Fuermaier et al., 2018). For instance, impulsive behavior can also arise in non-reward-related settings, such as when perceptually discriminating size differences where less accurate and faster responses have been observed in people with addiction disorders (Banca et al., 2016; Lawrence et al., 2009; Paasche et al., 2019) and borderline personality disorder (Berlin and Rolls, 2004). Perceptual deficiencies are also prevalent in clinical populations with attention-deficit hyperactivity disorder (ADHD) or Parkinson's disease, and thought to be linked to impairments in the dopaminergic system (Fuermaier et al., 2018; Sebastian et al., 2014; Caspers et al., 2017; Djamshidian et al., 2014; Herz et al., 2014; Pote et al., 2016). Causal evidence from deep-brain stimulation (DBS), in particular, has shown that disrupting the STN lowered decision thresholds, thus increasing this perceptual impulsivity among Parkinson's patients (Frank, 2007; Herz et al., 2016, 2017).

Our findings that TMS of the left SFS causally and selectively lowered the decision boundary during perceptual decisions suggest that the SFS may also be involved in speed-accuracy tradeoff regulation, suggesting that the lateral prefrontal cortex may be functionally integrated with these cortico-striatal and cortico-subthalamic nuclei (STN) pathways (Bogacz et al., 2010; Forstmann et al., 2010; Green et al., 2012, 2013). Overall, impulsive behavior is not exclusive to the reward domain, and our results suggest that there is something to gain from understanding impulsive behavior in non-reward settings requiring decisions on perceptual information. Maladaptive behavior may not only reflect individual wants or likings, often assumed by addiction studies, but could also be a function of low-level sensory or higher-order cognitive processes that have so far not fully been accounted for (Bartoshuk et al., 2006; Demmel and Schrenk, 2003; Fuermaier et al., 2018). This may have serious implications for how cognitive therapies or interventions are designed, and our findings may provide useful insights in guiding such future work. Particularly, it is worth exploring to what degree the left SFS and its connections are structurally or functionally different in clinical populations, and whether these impulsive tendencies can be captured by sequential sampling models, such as the DDM.

REFERENCES:

Ahmad, J., Swan, G., Bowman, H., Wyble, B., Nobre, A.C., Shapiro, K.L., and McNab, F. (2017). Competitive interactions affect working memory performance for both simultaneous and sequential stimulus presentation. Scientific Reports *7*, 4785.

Armstrong, K.M., and Moore, T. (2007). Rapid enhancement of visual cortical response discriminability by microstimulation of the frontal eye field. Proceedings of the National Academy of Sciences *104*.

Armstrong, K.M., Chang, M.H., and Moore, T. (2009). Selection and Maintenance of Spatial Information by Frontal Eye Field Neurons. Journal of Neuroscience *29*.

Awh, E., and Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. Trends in Cognitive Sciences *5*, 119–126.

Awh, E., Vogel, E.K., and Oh, S.-H. (2006). Interactions between attention and working memory. Neuroscience *139*, 201–208.

Badre, D., and D'Esposito, M. (2007). Functional Magnetic Resonance Imaging Evidence for a Hierarchical Organization of the Prefrontal Cortex. Journal of Cognitive Neuroscience *19*, 2082–2099.

Balleine BW. (2007) The neural basis of choice and decision making. *Journal of Neuroscience* 27: 8159-60

Ballesta, S., and Padoa-Schioppa, C. (2019). Economic Decisions through Circuit Inhibition. Current Biology *29*, 3814-3824.e5.

Banca, P., Lange, I., Worbe, Y., Howell, N.A., Irvine, M., Harrison, N.A., Moutoussis, M., and Voon, V. (2016). Reflection impulsivity in binge drinking: Behavioural and volumetric correlates. Addiction Biology *21*, 504–515.

Barack, D.L., and Platt, M.L. (2017). Engaging and Exploring: Cortical Circuits for Adaptive Foraging Decisions.

Barbas, H., and Mesulam, M.-M. (1981). Organization of afferent input to subdivisions of area 8 in the rhesus monkey. The Journal of Comparative Neurology *200*, 407–431.

Barbey, A.K., Koenigs, M., and Grafman, J. (2013). Dorsolateral prefrontal contributions to human working memory. Cortex *49*, 1195–1205.

Bartoshuk, L.M., Duffy, V.B., Hayes, J.E., Moskowitz, H.R., and Snyder, D.J. (2006). Psychophysics of sweet and fat perception in obesity: problems, solutions and new perspectives. Philosophical Transactions of the Royal Society B: Biological Sciences *361*.

Basten, U., Biele, G., Heekeren, H.R., and Fiebach, C.J. (2010). How the brain integrates costs and benefits during decision making. Proceedings of the National Academy of Sciences *107*, 21767–21772.

Beck, D.M., and Kastner, S. (2007). Stimulus similarity modulates competitive interactions in human visual cortex. Journal of Vision *7*, 19.

Beck, A., Schlagenhauf, F., Wüstenberg, T., Hein, J., Kienast, T., Kahnt, T., Schmack, K., Hägele, C., Knutson, B., Heinz, A., et al. (2009). Ventral striatal activation during reward anticipation correlates with impulsivity in alcoholics. Biological Psychiatry *66*, 734–742.

Berlin, H.A., and Rolls, E.T. (2004). Time Perception, Impulsivity, Emotionality, and Personality in Self-Harming Borderline Personality Disorder Patients. Journal of Personality Disorders *18*, 358–378.

Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD. (2006) The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review* 113: 700-65

Bogacz, R., Wagenmakers, E.J., Forstmann, B.U., and Nieuwenhuis, S. (2010). The neural basis of the speed-accuracy tradeoff. Trends in Neurosciences *33*, 10–16.

Brody, C.D., and Hanks, T.D. (2016). Neural underpinnings of the evidence accumulator. Current Opinion in Neurobiology *37*, 149–157.

Brunton, B.W., Botvinick, M.M., and Brody, C.D. (2013). Rats and Humans Can Optimally Accumulate Evidence for Decision-Making. Science *340*, 95–98.

Bullier, J., Schall, J.D., and Morel, A. (1996). Functional streams in occipito-frontal connections in the monkey. Behavioural Brain Research *76*, 89–97.

Busemeyer JR, Townsend JT. (1993). Decision Field-Theory - a Dynamic Cognitive Approach to Decision-Making in an Uncertain Environment. *Psychological Review* 100: 432-59

Carandini, M. (2012). From circuits to behavior: a bridge too far? Nature Neuroscience *15*, 507–509.

Caspers, J., Mathys, C., Hoffstaedter, F., Südmeyer, M., Cieslik, E.C., Rubbert, C., Hartmann, C.J., Eickhoff, C.R., Reetz, K., Grefkes, C., et al. (2017). Differential Functional Connectivity Alterations of Two Subdivisions within the Right dlPFC in Parkinson's Disease. Frontiers in Human Neuroscience *11*.

Cavanagh, J.F., Wiecki, T. v, Cohen, M.X., Figueroa, C.M., Samanta, J., Sherman, S.J., and Frank, M.J. (2011). Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. Nature Neuroscience *14*, 1462–1467.

Chawla M, Miyapuram KP. (2018). Context-Sensitive Computational Mechanisms of Decision Making. *J Exp Neurosci* 12: 1179069518809057

Courtney, S.M., Petit, L., Haxby, J. v., and Ungerleider, L.G. (1998). The role of prefrontal cortex in working memory: examining the contents of consciousness. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences *353*, 1819–1828.

Deco, G., Rolls, E.T., Albantakis, L., and Romo, R. (2013). Brain mechanisms for perceptual and reward-related decision-making. Progress in Neurobiology *103*, 194–213.

Demmel, R., and Schrenk, J. (2003). Sensory evaluation of alcohol-related and neutral stimuli. Addictive Behaviors *28*.

de Martino, B., Fleming, S.M., Garrett, N., and Dolan, R.J. (2013). Confidence in value-based choice. Nature Neuroscience *16*, 105–110.

di Lazzaro, V., Pilato, F., Saturno, E., Oliviero, A., Dileone, M., Mazzone, P., Insola, A., Tonali, P.A., Ranieri, F., Huang, Y.Z., et al. (2005). Theta-burst repetitive transcranial magnetic stimulation suppresses specific excitatory circuits in the human motor cortex. The Journal of Physiology *565*, 945–950.

di Lazzaro, V., Ziemann, U., and Lemon, R.N. (2008). State of the art: Physiology of transcranial motor cortex stimulation. Brain Stimulation *1*, 345–362.

Ding, L., and Gold, J.I. (2012a). Neural Correlates of Perceptual Decision Making before, during, and after Decision Commitment in Monkey Frontal Eye Field. Cerebral Cortex *22*, 1052–1067.

Ding, L., and Gold, J.I. (2012b). Separate, Causal Roles of the Caudate in Saccadic Choice and Execution in a Perceptual Decision Task. Neuron *75*, 865–874.

Ditz, H.M., and Nieder, A. (2020). Format-dependent and format-independent representation of sequential and simultaneous numerosity in the crow endbrain. Nature Communications *11*, 686.

Dixon ML, De La Vega A, Mills C, Andrews-Hanna J, Spreng RN, et al. (2018). Heterogeneity within the frontoparietal control network and its relationship to the default and dorsal attention networks. *Proc Natl Acad Sci U S A* 115: E1598-E607

Djamshidian, A., O'Sullivan, S.S., Lawrence, A.D., Foltynie, T., Aviles-Olmos, I., Magdalinou, N., Tomassini, A., Warner, T.T., Lees, A.J., and Averbeck, B.B. (2014). Perceptual decision-making in patients with Parkinson's disease. Journal of Psychopharmacology *28*, 1149–1154.

Dodds CM, Morein-Zamir S, Robbins TW. (2011). Dissociating inhibition, attention, and response control in the frontoparietal network using functional magnetic resonance imaging. *Cereb Cortex* 21: 1155-65

Domenech, P., and Dreher, J.-C. (2010). Decision Threshold Modulation in the Human Brain. Journal of Neuroscience *30*, 14305–14317.

Domenech, P., and Koechlin, E. (2015). Executive control and decision-making in the prefrontal cortex. Current Opinion in Behavioral Sciences *1*, 101–106.

Domenech, P., Redouté, J., Koechlin, E., and Dreher, J.-C. (2017). The Neuro-Computational Architecture of Value-based Selection in the Human Brain. Cerebral Cortex 1–17.

Dutilh G, Rieskamp J. (2016). Comparing perceptual and preferential decision making. *Psychon Bull Rev* 23: 723-37

Duncan, J., and Owen, A.M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. Trends in Neurosciences *23*, 475–483.

Erlich, J.C., Brunton, B.W., Duan, C.A., Hanks, T.D., and Brody, C.D. (2015). Distinct effects of prefrontal and parietal cortex inactivations on an accumulation of evidence task in the rat. ELife *4*.

Feltgen, Q., and Daunizeau, J. (2020). Fitting drift-diffusion decision models to trial-by-trial data. BioRxiv.

Forstmann, B.U., and Wagenmakers, E.-J. (2015). An introduction to model-based cognitive neuroscience (Springer).

Filimon, F., Philiastides, M.G., Nelson, J.D., Kloosterman, N.A., and Heekeren, H.R. (2013). How Embodied Is Perceptual Decision Making? Evidence for Separate Processing of Perceptual and Motor Decisions. Journal of Neuroscience *33*, 2121–2136.

Forstmann, B.U., Wagenmakers, E.-J., Eichele, T., Brown, S., and Serences, J.T. (2011). Reciprocal relations between cognitive neuroscience and formal cognitive models: opposites attract? Trends in Cognitive Sciences *15*, 272–279.

Forstmann, B.U., Anwander, A., Schafer, A., Neumann, J., Brown, S., Wagenmakers, E.-J., Bogacz, R., and Turner, R. (2010). Cortico-striatal connections predict control over speed and accuracy in

perceptual decision making. Proceedings of the National Academy of Sciences *107*, 15916–15920.

Frank, M.J. (2007). Medication in Parkinsonism. Science *1309*, 1309–1312.

Fuermaier, A.B.M., Hüpen, P., de Vries, S.M., Müller, M., Kok, F.M., Koerts, J., Heutink, J., Tucha, L., Gerlach, M., and Tucha, O. (2018). Perception in attention deficit hyperactivity disorder. ADHD Attention Deficit and Hyperactivity Disorders *10*, 21–47

Gatto, E.M., and Aldinio, V. (2019). Impulse Control Disorders in Parkinson's Disease. A Brief and Comprehensive Review. Frontiers in Neurology *10*.

Gazzaley, A., Rissman, J., and D'Esposito, M. (2004). Functional connectivity during working memory maintenance. Cognitive, Affective, & Behavioral Neuroscience *4*, 580–599.

Gazzaley, A., Rissman, J., Cooney, J., Rutman, A., Seibert, T., Clapp, W., and D'Esposito, M. (2007). Functional Interactions between Prefrontal and Visual Association Cortex Contribute to Top-Down Modulation of Visual Processing. Cerebral Cortex *17*, i125–i135.

Georgiev, D., Rocchi, L., Tocco, P., Speekenbrink, M., Rothwell, J.C., and Jahanshahi, M. (2016). Continuous Theta Burst Stimulation Over the Dorsolateral Prefrontal Cortex and the Pre-SMA Alter Drift Rate and Response Thresholds Respectively During Perceptual Decision-Making. Brain Stimulation *9*, 601–608.

Gerstner W, Sprekeler H, Deco G. (2012). Theory and Simulation in Neuroscience. *Science* 338: 60-65.

Glickman, M., Sharoni, O., Levy, D.J., Niebur, E., Stuphorn, V., and Usher, M. (2019). The formation of preference in risky choice. PLoS Computational Biology *15*, e1007201.

Glimcher, P.W., Kable, J., and Louie, K. (2007). Neuroeconomic Studies of Impulsivity: Now or Just as Soon as Possible? American Economic Review *97*, 142–147.

Gold JI, Shadlen MN. (2007). The neural basis of decision making. *Annual Review of Neuroscience* 30: 535-74

Goldman, P.S., and Nauta, W.J.H. (1977). Columnar distribution of cortico-cortical fibers in the frontal association, limbic, and motor cortex of the developing rhesus monkey. Brain Research *122*, 393–413.

Goldman-Rakic, P.S. (1987). Circuitry of the prefrontal cortex and the regulation of behavior by representational memory. In Handbook of Neurobiology, V.B. Mountcastle, ed. (Bethesda, MD: American Physiological Society), pp. 373–417.

Goldman-Rakic, P.S. (1984). Modular organization of prefrontal cortex. Trends in Neurosciences *7*, 419–424.

Goldman-Rakic, P.S. (1994). The Issue of Memory in the Study of Prefrontal Function.

Goldman-Rakic, P.S. (1995). Architecture of the Prefrontal Cortex and the Central Executive. Annals of the New York Academy of Sciences *769*, 71–84.

Goldman-Rakic, P.S., and Leung, H.-C. (2002). Functional Architecture of the Dorsolateral Prefrontal Cortex in Monkeys and Humans (Oxford University Press).

Green, N., Biele, G.P., and Heekeren, H.R. (2012). Changes in Neural Connectivity Underlie Decision Threshold Modulation for Reward Maximization. Journal of Neuroscience *32*, 14942–14950.

Green, N., Bogacz, R., Huebl, J., Beyer, A.K., Kühn, A.A., and Heekeren, H.R. (2013). Reduction of influence of task difficulty on perceptual decision making by stn deep brain stimulation. Current Biology *23*, 1681–1684.

Grueschow, M., Polania, R., Hare, T.A., and Ruff, C.C. (2015). Automatic versus Choice-Dependent Value Representations in the Human Brain. Neuron *85*, 874–885.

Grueschow, M., Polania, R., Hare, T.A., and Ruff, C.C. (2018). Arousal Optimizes Neural Evidence Representation for Human Decision-Making. SSRN Electronic Journal.

Grueschow, M., Kleim, B., and Ruff, C.C. (2020). Role of the locus coeruleus arousal system in cognitive control. Journal of Neuroendocrinology *32*.

Hanks, T.D., and Summerfield, C. (2017). Perceptual Decision Making in Rodents, Monkeys, and Humans. Neuron *93*, 15–31.

Hanks, T.D., Kopec, C.D., Brunton, B.W., Duan, C.A., Erlich, J.C., and Brody, C.D. (2015). Distinct relationships of parietal and prefrontal cortices to evidence accumulation. Nature *520*, 220–223.

Hare, T.A., Camerer, C.F., and Rangel, A. (2009). Self-Control in Decision-Making Involves Modulation of the vmPFC Valuation System. Science *324*, 646–648.

Hare, T.A., Schultz, W., Camerer, C.F., O'Doherty, J.P., and Rangel, A. (2011). Transformation of stimulus value signals into motor commands during simple choice. Proceedings of the National Academy of Sciences *108*, 18120–18125.

Heekeren, H.R., Marrett, S., Bandettini, P.A., and Ungerleider, L.G. (2004). A general mechanism for perceptual decision-making in the human brain. Nature *431*, 859–862.

Heekeren, H.R., Marrett, S., Ruff, D.A., Bandettini, P.A., and Ungerleider, L.G. (2006). Involvement of human left dorsolateral prefrontal cortex in perceptual decision making is independent of response modality. Proceedings of the National Academy of Sciences *103*, 10023–10028.

Heekeren, H.R., Marrett, S., and Ungerleider, L.G. (2008). The neural systems that mediate human perceptual decision making. Nature Reviews Neuroscience *9*, 467–479.

Heitz RP, Schall JD. (2012). Neural mechanisms of speed-accuracy tradeoff. *Neuron* 76: 616-28

Herz, D.M., Eickhoff, S.B., Løkkegaard, A., and Siebner, H.R. (2014). Functional neuroimaging of motor control in parkinson's disease: A meta-analysis. Human Brain Mapping *35*, 3227–3237.

Herz, D.M., Zavala, B.A., Bogacz, R., and Brown, P. (2016). Neural Correlates of Decision Thresholds in the Human Subthalamic Nucleus. Current Biology *26*, 916–920.

Herz, D.M., Tan, H., Brittain, J.S., Fischer, P., Cheeran, B., Green, A.L., Fitzgerald, J., Aziz, T.Z., Ashkan, K., Little, S., et al. (2017). Distinct mechanisms mediate speed-accuracy adjustments in cortico-subthalamic networks. ELife *6*, 1–25.

Heyes, S.B., Adam, R.J., Urner, M., van der Leer, L., Bahrami, B., Bays, P.M., and Husain, M. (2012). Impulsivity and rapid decision-making for reward. Frontiers in Psychology *3*, 1–11.

Huang, Y.Z., Edwards, M.J., Rounis, E., Bhatia, K.P., and Rothwell, J.C. (2005). Theta burst stimulation of the human motor cortex. Neuron *45*, 201–206.

Hutcherson, C.A., Bushong, B., and Rangel, A. (2015). A Neurocomputational Model of Altruistic Choice and Its Implications. Neuron *87*, 451–462.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. Journal of Experimental Psychology *45*, 188.

Ide, J.S., Tung, H.C., Yang, C.-T., Tseng, Y.-C., and Li, C.-S.R. (2017). Barratt Impulsivity in Healthy Adults Is Associated with Higher Gray Matter Concentration in the Parietal Occipital Cortex that Represents Peripheral Visual Field. Frontiers in Human Neuroscience *11*, 1–11.

Jao Keehn, R.J., Nair, S., Pueschel, E.B., Linke, A.C., Fishman, I., and Müller, R.-A. (2019). Atypical local and distal patterns of occipito-frontal functional connectivity are related to symptom severity in autism. Cerebral Cortex *29*, 3319–3330.

Kayser AS, Buchsbaum BR, Erickson DT, D'Esposito M. (2010). The Functional Anatomy of a Perceptual Decision in the Human Brain. *J Neurophysiol* 103: 1179-94

Keuken MC, Muller-Axt C, Langner R, Eickhoff SB, Forstmann BU, Neumann J. (2014). Brain networks of perceptual decision-making: an fMRI ALE meta-analysis. *Frontiers in Human Neuroscience* 8.

Kim JN, Shadlen MN. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nature Neuroscience* 2: 176-85

Koechlin, E. (2003). The Architecture of Cognitive Control in the Human Prefrontal Cortex. Science *302*, 1181–1185.

Koechlin, E., and Jubault, T. (2006). Broca's Area and the Hierarchical Organization of Human Behavior. Neuron *50*, 963–974.

Koechlin, E., and Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. Trends in Cognitive Sciences *11*, 229–235.

Koechlin, E., Basso, G., Pietrini, P., Panzer, S., and Grafman, J. (1999). The role of the anterior prefrontal cortex in human cognition. Nature *399*, 148–151.

Krajbich I. (2019). Accounting for attention in sequential sampling models of decision making. *Curr Opin Psychol* 29: 6-11

Krawczyk DC. (2002). Contributions of the prefrontal cortex to the neural basis of human decision making. *Neurosci Biobehav R* 26: 631-64

Lawrence, A.J., Luty, J., Bogdan, N.A., Sahakian, B.J., and Clark, L. (2009). Impulsivity and response inhibition in alcohol dependence and problem gambling. Psychopharmacology *207*, 163–172.

Levy, R., and Goldman-Rakic, P.S. (2000). Segregation of working memory functions within the dorsolateral prefrontal cortex. In Executive Control and the Frontal Lobe: Current Issues, (Berlin, Heidelberg: Springer Berlin Heidelberg),

Li, C.R., Luo, X., Yan, P., Bergquist, K., and Sinha, R. (2009). Altered Impulse Control in Alcohol Dependence: Neural Measures of Stop Signal Performance. Alcoholism: Clinical and Experimental Research *33*, 740–750.

Liu, T., and Pleskac, T.J. (2011). Neural correlates of evidence accumulation in a perceptual decision task. *48824*, 2383–2398.

Logothetis, N.K. (2008). What we can do and what we cannot do with fMRI. Nature *453*, 869–878.

Luck, S.J., Chelazzi, L., Hillyard, S.A., and Desimone, R. (1997). Neural Mechanisms of Spatial Selective Attention in Areas V1, V2, and V4 of Macaque Visual Cortex. Journal of Neurophysiology *77*, 24–42.

Maier, S.U., Makwana, A.B., and Hare, T.A. (2015). Acute Stress Impairs Self-Control in Goal-Directed Choice by Altering Multiple Functional Connections within the Brain's Decision Circuits. Neuron *87*, 621–631.

Mante V, Sussillo D, Shenoy KV, Newsome WT. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* 503: 78-+

Mäntylä, T., Still, J., Gullberg, S., and del Missier, F. (2012). Decision Making in Adults With ADHD. Journal of Attention Disorders *16*, 164–173.

Marek S, Dosenbach NUF. (2018). The frontoparietal network: function, electrophysiology, and importance of individual precision mapping. *Dialogues Clin Neurosci* 20: 133-40

Marr, D. (2010). Vision: A computational investigation into the human representation and processing of visual information (MIT press).

Mawase, F., Lopez, D., Celnik, P.A., and Haith, A.M. (2018). Movement repetition facilitates response preparation. Cell Reports *24*, 801–808.

Mayer, A.R., Ryman, S.G., Hanlon, F.M., Dodd, A.B., and Ling, J.M. (2016). Look Hear! The Prefrontal Cortex is Stratified by Modality of Sensory Input During Multisensory Cognitive Control. Cerebral Cortex.

Mazurek ME, Roitman JD, Ditterich J, Shadlen MN. (2003). A role for neural integrators in perceptual decision making. *Cereb Cortex* 13: 1257-69

Mcclure, S.M., and Bickel, W.K. (2014). A dual-systems perspective on addiction: Contributions from neuroimaging and cognitive training. Annals of the New York Academy of Sciences *1327*, 62–78.

Meng, Y., Hu, X., Bachevalier, J., and Zhang, X. (2016). Decreased functional connectivity in dorsolateral prefrontal cortical networks in adult macaques with neonatal hippocampal lesions: Relations to visual working memory deficits. Neurobiology of Learning and Memory *134*, 31–37.

Mischel, W., Ebbesen, E.B., and Raskoff Zeiss, A. (1972). Cognitive and attentional mechanisms in delay of gratification. Journal of Personality and Social Psychology *21*, 204–218.

Moeller, F.G., Barratt, E.S., Ph, D., Dougherty, D.M., Schmitz, J.M., and Swann, A.C. (2001). Reviews and Overviews Psychiatric Aspects of Impulsivity. Psychiatry Interpersonal and Biological Processes *158*, 1783–1793.

Mogilner, C., Shiv, B., and Iyengar, S.S. (2013). Eternal Quest for the Best: Sequential (vs. Simultaneous) Option Presentation Undermines Choice Commitment. Journal of Consumer Research *39*, 1300–1312.

Moran R. (2015). Optimal decision making in heterogeneous and biased environments. *Psychon Bull Rev* 22: 38-53

Mowbray, G.H., and Rhoades, M. v (1959). On the reduction of choice reaction times with practice. Quarterly Journal of Experimental Psychology *11*, 16–23.

Muhle-Karbe PS, Jiang J, Egner T. (2018). Causal Evidence for Learning-Dependent Frontal Lobe Contributions to Cognitive Control. *Journal of Neuroscience* 38: 962-73

Mulder MJ, van Maanen L, Forstmann BU. (2014). Perceptual decision neurosciences - a model-based review. *Neuroscience* 277: 872-84

Murd, C., Moisa, M., Grueschow, M., Polania, R., and Ruff, C.C. (2020). Causal contributions of human frontal eye fields to distinct aspects of decision formation. Scientific Reports *10*, 7317.

Nee, D.E., and D'Esposito, M. (2016). The hierarchical organization of the lateral prefrontal cortex. ELife *5*.

Owen, A.M. (1997). The Functional Organization of Working Memory Processes Within Human Lateral Frontal Cortex: The Contribution of Functional Neuroimaging. European Journal of Neuroscience *9*, 1329–1339.

Owen, A.M., Evans, A.C., and Petrides, M. (1996). Evidence for a Two-Stage Model of Spatial Working Memory Processing within the Lateral Frontal Cortex: A Positron Emission Tomography Study. Cerebral Cortex *6*, 31–38.

Paasche, C., Weibel, S., Wittmann, M., and Lalanne, L. (2019). Time perception and impulsivity: A proposed relationship in addictive disorders. Neuroscience & Biobehavioral Reviews *106*, 182–201.

Palmeri, T.J., Love, B.C., and Turner, B.M. (2017). Model-based cognitive neuroscience. Journal of Mathematical Psychology *76*.

Peters, J., and D'Esposito, M. (2020). The drift diffusion model as the choice rule in inter-temporal and risky choice: a case study in medial orbitofrontal cortex lesion patients and controls. PLoS Computational Biology *16*, e1007615.

Petrides, M. (1995). Functional Organization of the Human Frontal Cortex for Mnemonic Processing.. Annals of the New York Academy of Sciences *769*, 85–96.

Petrides, M. (2005). Lateral prefrontal cortex: architectonic and functional organization. Philosophical Transactions of the Royal Society B: Biological Sciences *360*, 781–795.

Philiastides, M.G., Auksztulewicz, R., Heekeren, H.R., and Blankenburg, F. (2011). Causal role of dorsolateral prefrontal cortex in human perceptual decision making. Current Biology *21*, 980–983.

Piet AT, Erlich JC, Kopec CD, Brody CD. (2017). Rat Prefrontal Cortex Inactivations during Decision Making Are Explained by Bistable Attractor Dynamics. *Neural Comput* 29: 2861-86

Ploran EJ, Nelson SM, Velanova K, Donaldson DI, Petersen SE, Wheeler ME. (2007). Evidence accumulation and the moment of recognition: Dissociating perceptual recognition processes using fMRI. *Journal of Neuroscience* 27: 11912-24

Polanía, R., Krajbich, I., Grueschow, M., and Ruff, C.C. (2014). Neural Oscillations and Synchronization Differentially Support Evidence Accumulation in Perceptual and Value-based Decision Making. Neuron *82*, 709–720.

Polania R, Moisa M, Opitz A, Grueschow M, Ruff CC. (2015). The precision of value-based choices depends causally on fronto-parietal phase coupling. *Nat Commun* 6

Poldrack, R. (2006). Can cognitive processes be inferred from neuroimaging data? Trends in Cognitive Sciences *10*, 59–63.

Postle, B.R. (2016). How Does the Brain Keep Information "in Mind"? Current Directions in Psychological Science *25*, 151–156.

Postle, B.R., Berger, J.S., Taich, A.M., and D'Esposito, M. (2000). Activity in Human Frontal Cortex Associated with Spatial Working Memory and Saccadic Behavior. Journal of Cognitive Neuroscience *12*, 2–14.

Pote, I., Torkamani, M., Kefalopoulou, Z.M., Zrinzo, L., Limousin-Dowsey, P., Foltynie, T., Speekenbrink, M., and Jahanshahi, M. (2016). Subthalamic nucleus deep brain stimulation induces impulsive action when patients with Parkinson's disease act under speed pressure. Experimental Brain Research *234*, 1837–1848.

Purcell BA, Heitz RP, Cohen JY, Schall JD, Logan GD, Palmeri TJ. (2010). Neurally Constrained Modeling of Perceptual Decision Making. *Psychological Review* 117: 1113-43

Rahnev, D., Nee, D.E., Riddle, J., Larson, A.S., and D'Esposito, M. (2016). Causal evidence for frontal cortex organization for perceptual decision making. Proceedings of the National Academy of Sciences *113*, 6059–6064.

Ramsey, J.D., Hanson, S.J., Hanson, C., Halchenko, Y.O., Poldrack, R.A., and Glymour, C. (2010). Six problems for causal inference from fMRI. NeuroImage *49*, 1545–1558.

Raos, V., and Savaki, H.E. (2016). The Role of the Prefrontal Cortex in Action Perception. Cerebral Cortex *27*, 4677–4690.

Ratcliff, R. (1978). A theory of memory retrieval. Psychological Review *85*, 59–108.

Ratcliff R, Mckoon G. (1988). A Retrieval Theory of Priming in Memory. *Psychological Review* 95: 385-408

Ratcliff R, Rouder JN. (1998). Modeling response times for two-choice decisions. *Psychol Sci* 9: 347-56

Ratcliff, R., and McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. Neural Computation *20*, 873–922.

Ratcliff, R., Voskuilen, C., and McKoon, G. (2018). Internal and external sources of variability in perceptual decision-making. Psychological Review *125*, 33–46.

Ricker, T.J., and Cowan, N. (2014). Differences between presentation methods in working memory procedures: A matter of working memory consolidation. Journal of Experimental Psychology: Learning, Memory, and Cognition *40*, 417–428.

Rowe, J.B., Toni, I., Josephs, O., Frackowiak, R.S.J., and Passingham, R.E. (2000). The Prefrontal Cortex: Response Selection or Maintenance Within Working Memory? Science *288*, 1656–1660.

Rudorf, S., and Hare, T.A. (2014). Interactions between Dorsolateral and Ventromedial Prefrontal Cortex Underlie Context-Dependent Stimulus Valuation in Goal-Directed Choice. Journal of Neuroscience *34*, 15988–15996.

Ruff, C.C., Blankenburg, F., Bjoertomt, O., Bestmann, S., Freeman, E., Haynes, J.-D., Rees, G., Josephs, O., Deichmann, R., and Driver, J. (2006). Concurrent TMS-fMRI and Psychophysics Reveal Frontal Influences on Human Retinotopic Visual Cortex. Current Biology *16*, 1479–1488.

Scolari M, Seidl-Rathkopf KN, Kastner S. (2015). Functions of the human frontoparietal attention network: Evidence from neuroimaging. *Curr Opin Behav Sci* 1: 32-39

Sebastian, A., Jung, P., Krause-Utz, A., Lieb, K., Schmahl, C., and TÃ¼scher, O. (2014). Frontal Dysfunctions of Impulse Control â€“ A Systematic Review in Borderline Personality Disorder and Attention-Deficit/Hyperactivity Disorder. Frontiers in Human Neuroscience *8*.

Serences, J.T., Ester, E.F., Vogel, E.K., and Awh, E. (2009). Stimulus-Specific Delay Activity in Human Primary Visual Cortex. Psychological Science *20*, 207–214.

Shadlen, M.N., and Newsome, W.T. (1996). Motion perception: seeing and deciding. Proceedings of the National Academy of Sciences *93*, 628–633.

Skagerlund, K., Karlsson, T., and Träff, U. (2016). Magnitude Processing in the Brain: An fMRI Study of Time, Space, and Numerosity as a Shared Cortical System. Frontiers in Human Neuroscience *10*, 1–12.

Sokol-Hessner, P., Hutcherson, C., Hare, T., and Rangel, A. (2012). Decision value computation in DLPFC and VMPFC adjusts to the available decision time. European Journal of Neuroscience *35*, 1065–1074.

Smith PL, Ratcliff R. (2004). Psychology and neurobiology of simple decisions. *Trends Neurosci* 27: 161-68

Snow, P. (2016) The structural and functional organization of cognition. *Frontiers in Human Neurosicence.*

Specker, S., Carlson, G., Christenson, G., and Marcotte, M. (1995). Impulse Control Disorders and Attention Deficit Disorder in Pathological Gamblers. Annals of Clinical Psychiatry *7*, 175–179.

Starns, J.J., and Ma, Q. (2018). Response biases in simple decision making: Faster decision making, faster response execution, or both? Psychonomic Bulletin & Review *25*, 1535–1541.

Stewart, N., Hermens, F., and Matthews, W.J. (2016). Eye movements in risky choice. Journal of Behavioral Decision Making *29*, 116–136.

Summerfield C, Tsetsos K. (2012). Building bridges between perceptual and economic decision-making: neural and computational mechanism. *Front Neurosci-Switz* 6

Tajima S, Drugowitsch J, Pouget A. (2016). Optimal policy for value-based decision-making. *Nat Commun* 7: 12400

Tosoni, A., Galati, G., Romani, G.L., and Corbetta, M. (2008). Sensory-motor mechanisms in human parietal cortex underlie arbitrary visual decisions. Nature Neuroscience *11*, 1446–1453.

Turner, B.M., van Maanen, L., and Forstmann, B.U. (2015). Informing cognitive abstractions through neuroimaging: the neural drift diffusion model. Psychological Review *122*, 312–336.

Usher M, McClelland JL. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review* 108: 550-92

Vandekerckhove, J., Tuerlinckx, F., and Lee, M.D. (2011). Hierarchical diffusion models for two-choice response times. Psychological Methods *16*, 44–62.

White, C.N., Servant, M., and Logan, G.D. (2018). Testing the validity of conflict drift-diffusion models for use in estimating cognitive processes: A parameter-recovery study. Psychonomic Bulletin & Review *25*, 286–301.

White CN, Mumford JA, Poldrack RA. (2012). Perceptual Criteria in the Human Brain. *Journal of Neuroscience* 32: 16716-24

Wycoco, V., Shroff, M., Sudhakar, S., and Lee, W. (2013). White Matter Anatomy. Neuroimaging Clinics of North America *23*.

Wijeakumar, S., Ambrose, J.P., Spencer, J.P., and Curtu, R. (2017). Model-based functional neuroimaging using dynamic neural fields: An integrative cognitive neuroscience approach. Journal of Mathematical Psychology *76*, 212–235.

Zanto, T.P., Rubens, M.T., Thangavel, A., and Gazzaley, A. (2011). Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. Nature Neuroscience *14*, 656–661.

Zhang, R., Hu, Z., Debi, R., Zhang, L., Li, H., and Liu, Q. (2013). Neural Processes Underlying the“Same”-“Different” Judgment of Two Simultaneously Presented Objects- An EEG Study. PLoS ONE *8*, e81737.

MATERIALS AND METHODS

Participants. Twenty healthy right-handed volunteers (ages 20-30; 8 female) with normal or corrected-to-normal vision participated in the study. Participants were fully informed about the features of the experiment. No participant suffered from any neurological or psychological disorder or took medication that interfered with their participation in the study. Participants received monetary compensation for participation and performance of the perceptual choices, as well as a one food item to consume after the experiment depending on a random value-based choice trial. The experiments conformed to the Declaration of Helsinki and the experimental protocol was approved by the Ethics Committee of the Canton of Zurich.

Experimental Paradigm. We asked participants to refrain from eating for 3 hours before the start of the experiment. Our experiments took place between 0800 and 1900 hr during the day. The experiment consisted of two steps: (1) a rating task outside the scanner and (2) a decision-making task inside the scanner. During the rating task, we asked participants to provide perceptual- and value-based ratings of the same set of 61 food images using an on-screen slider scale. All of the food items were in stock in our lab and participants were informed about this via visual inspection. For perceptual ratings, participants rated—on a scale from 5 to 100 percent in steps of 5 percent— how much of the black background within the white square perimeter was occupied by the food item. For value-based ratings, participants rated—on a scale from 5 to 100 in steps of 5—how much they wanted to eat the presented food item at the end of the experiment. We instructed participants that the

midpoint of the scale in value-based ratings indicated indifference. The ratings phase required participants to rate the same food items twice for each task.

After rating the food items, an algorithm selected a balanced set of perceptual and value-based trials divided into four evidence levels, E. The evidence levels are based on the absolute difference between the average ratings of the food items paired in each trial. We define perceptual evidence as the absolute size difference between the two food items. On the other hand, we define value-based evidence as the absolute value difference between the two food items. In particular, the evidence levels for perceptual trials, E_p, are:

$$E_p = |r_{\mathbf{biggest}} - r_{\mathbf{smallest}}| \in \{5\%,10\%,15\%,20\%\} \quad (1)$$

while the evidence levels for value-based trials, E_v, are:

$$E_v = |r_{\mathbf{best}} - r_{\mathbf{worst}}| \in \{1,2,3,4\} \quad (2)$$

where $r = \frac{r_1 + r_2}{2}$ is the average food item from the two ratings while $r_{\mathbf{biggest}} - r_{\mathbf{smallest}}$ and $r_{\mathbf{best}} - r_{\mathbf{worst}}$ represent the ratings' difference for the pairs presented for perceptual and value-based choices, respectively.

Inside the scanner, participants performed the decision-making task for which they chose between two food items, based on whether they were accumulating perceptual or value-based evidence. We matched the visual sensory stimuli of the food items as well as their motor outputs across the two choice types. The only difference was the type of evidence participants had to accumulate to make a choice. Each trial started with presentation of a central fixation marker (length $\sim$ 0.8°, height $\sim$ 0.3°). Next, a centrally presented word indicated whether participants would perform a perceptual (word 'AREA') or value-based

(word 'LIKE') choice. On the subsequent screen, the task cue was replaced by either the letter '*A*' or '*L*' (~ 0.2°) to remind participants that they were in a perceptual or value-based block, respectively. Two food items were simultaneously displayed, one above and one below the screen (*y* eccentricity 3.6°; a white square of 6° width surrounded each food item).

Blocks alternated between perceptual and value-based choices in a given session (7-9 trials per task block). Participants pressed one of two buttons on a keypad with their right middle finger (upper item) or right index finger (lower item) to indicate their choice. On a given trial, participants had 3 seconds for their choice; otherwise, the trial would be regarded as a 'missed trial' and would not enter the analysis. After the experiment, participants stayed in the room with the experimenter while they ate the food that was selected based on the participants choice in one randomly selected VDM trial.

Participants made correct choices when they chose the food item with the higher rating as indicated in the double ratings task prior to entering the scanner. The experiment had a total of 256 trials divided into 8 sessions of 32 trials each. The first 4 sessions were pre-stimulation sessions where participants performed the task without stimulation. The last 4 sessions were post-stimulation sessions during which participants performed the choices with decreased neural excitability in the SFS due to the preceding continuous theta-burst stimulation. The 256 trials were fully balanced across all factors (trial type: perceptual or value-based; evidence levels: 1 to 4; correct response: up or down).

Stimulation Protocol. We applied continuous theta-burst stimulation (cTBS) (Huang et al., 2005; di Lazzaro et al., 2005, 2008) to exogenously induce cortical inhibition of our region of interest (ROI), an area in the left superior frontal sulcus (SFS) (MNI coordinates: $x = -24, y = 24, z = 36$) (Heekeren et al., 2004; Philiastides et al., 2011). Before the main fMRI experiment, we identified the stimulation site over the left SFS as well as each individual's stimulation intensity. In an initial fMRI session, we acquired high-resolution T1-

weighted 3D fast-field echo anatomical scans used for subsequent neuro-navigation (181 sagittal slices, matrix size = 256 × 256, voxel size = 1 mm^3, TR/TE/ TI = 8.3/2.26/181 ms, 3T Philips Achieva). The hand area of the left M1 (motor hotspot) was determined by identifying the first dorsal interosseous (FDI) movement-evoked potentials (MEPs) induced by transcranial magnetic stimulation (TMS) pulses. We delivered single monophasic TMS pulses using a figure-of-eight coil attached to the TMS stimulator. We then marked an equidistant circular grid on each individual's anatomical MRI scan using a neuro-navigation system over the hand motor region, located at the anterior portion of the central sulcus. We localized the optimal motor hotspot as the point in the grid that elicited the strongest FDI MEPs from TMS pulses. Once we selected the motor hotspot, we asked participants to activate their FDI by pressing their thumb and index finger at about 20% maximum force in order to obtain their active motor threshold (AMT). We defined the AMT as the minimal TMS intensity required to produce MEPs of $\geq$ 200 mV amplitude (measured with Magventure MRi-B91) in $\geq$ 5–10 consecutive pulses.

We retested the AMT by visually inspecting the FDI twitches triggered by TMS pulses over the marked optimal hotspot. The average AMT outside the scanner was 52.35 $\pm$ 6.27 percent while the AMT inside the scanner was 52.91 $\pm$ 6.18 percent. We applied cTBS at an intensity of 80% of the individual's AMT. The cTBS protocol contained bursts of 3 pulses at 50 Hz. This protocol has been shown to reduce cortical excitability for at least 30 minutes (Huang et al., 2005). Every burst was repeated at a rate of 5 Hz, resulting in 200 bursts with a total of 600 pulses delivered within 40 seconds.

Before moving our participant into the scanner, we marked the motor hotspot as well as the stimulation site on a swimming cap fixed in position by straps. Participants wore this cap while they were inside the scanner. Before the start of the fifth session, participants received cTBS over the left SFS. We used a figure-of-eight MR-compatible TMS coil (MRi-B91) attached to a TMS stimulator. After receiving stimulation, participants returned to the

scanner and proceeded to complete the last four sessions. On average, the post-stimulation fMRI task started 228 $\pm$ 41 sec after the end of theta-burst stimulation following established protocols from previous studies (Knecht et al., 2003; Philiastides et al., 2011; Thut and Pascual-Leone, 2010).

Given the established timeline of cTBS effects (Huang et al., 2005), we expected the stimulation effects to weaken over time due to neural recovery. Hence, we treated the first two post-stimulation sessions as the actual post-cTBS period and the last two post-stimulation sessions as the recovery period, in line with established procedures (Philiastides et al., 2011).

Differences-in-Differences Framework. We implemented a differences-in-differences (DID) framework to identify causal relationships based on stimulation-induced neural inhibition in the SFS. We used the identical DID contrast framework for behavioral, computational, neuroimaging and connectivity analyses. We employ the following notation: Task conditions $Task$ (perceptual ($Task = 1$) and value-based ($Task = 0$)) and stimulation conditions TMS (pre- ($TMS = 0$) and post-stimulation ($TMS = 1$)). Let V be our variable of interest, and either behavioral or neural. The causal treatment effect, $\boldsymbol{\phi}(V|Task,TMS)$, takes the following structural form:

$$\begin{aligned}\boldsymbol{\phi}(V|Task,TMS) = [\mathbb{E}(V|TMS = 1,Task = 1) - \mathbb{E}(V|TMS = 0,Task = 1)] \\ - [\mathbb{E}(V|TMS = 1,Task = 0) - \mathbb{E}(V|TMS = 0,Task = 0)]\end{aligned} \quad (3)$$

where $\mathbb{E}(V|Task,TMS)$ is the expected value of the variable of interest, V, given task, $Task$ and condition, TMS. The first difference on the right-hand side captures the average stimulation effect for perceptual choice while the second difference captures the average stimulation effect for value-based choice. The overall difference assumes that if behavior will be the same after stimulation, then $\phi = 0$ (Angrist and Pischke, 2009; Bertrand et al., 2004).

Informed by previous studies, we expect that left SFS stimulation will affect perceptual decisions (Heekeren et al., 2004, 2006; Philiastides et al., 2011) but not value-based decisions. If this holds, we expect the DID estimate to be negative, $\phi < 0$, suggesting inhibition of either behavior or neural activity.

Behavioral Analyses of Choice. We analyzed the influence of continuous theta-burst stimulation on choice using a logit regression on choices, ρ (correct = 1, incorrect = 0) over various regressors of interest, including stimulation condition, TMS (pre-cTBS = 0, post-cTBS = 1); task condition, $Task$ (perceptual = 1, value-based = 0); its interaction ($Task \times TMS$), which measures the causal stimulation effect, $\boldsymbol{\phi}$; and, other regressors, X^k that we used to control for during the regression analysis, and this includes the task-relevant evidence levels (*SD* for perceptual and *VD* for value-based, 1 to 4), response times (RTs), and task-irrelevant evidence levels (i.e. *VD* for perceptual and *SD* for value-based, 1 to 4). The full model is:

$$\Pr\left(\rho^{DID}_{t,c,s,i}\right) = \frac{1}{1 + \exp\left(-[\beta_0 + \beta_1 Task_{(t,c,s,i)} + \beta_2 TMS_{(t,c,s,i)} + \boldsymbol{\phi} Task_{(t,c,s,i)} TMS_{(t,c,s,i)} + \sum_{k=4}^{n} \beta_k X^k_{(t,c,s,i)}\right)} \quad (4)$$

where t indexes *Task, c* indexes *TMS*, s indexes subject, and i indexes trial. Since our model contains a DID interaction term, nonlinearity of the logit regression results in a non-zero estimate even if the true causal effect is zero, $\phi = 0$. To remove nonlinearity bias and isolate the true causal effect, we ran another logit regression without the interaction term,

$$\Pr\left(\rho^{NODID}_{t,c,s,i}\right) = \frac{1}{1 + \exp\left(-\left[\beta_0 + \beta_1 Task_{(t,c,s,i)} + \beta_2 TMS_{(t,c,s,i)} + \sum_{k=4}^{n} \beta_k X^k_{(t,c,s,i)}\right]\right)} \quad (5)$$

and we took the difference between the two logit models (Ai and Norton, 2003; Karaca-Mandic et al., 2012; Puhani, 2012),

$$\Pr\left(\rho_{(t,c,s,i)}^{TRUEDID}\right) = \Pr\left(\rho_{(t,c,s,i)}^{DID}\right) - \Pr\left(\rho_{(t,c,s,i)}^{NODID}\right) \quad (6)$$

We used cluster-robust standard errors at the subject level under the assumption that each individual performance is independent across participants. We also ran variations of the model to test for the effect's robustness, particularly GLMs with or without control variables, and using various stimulation runs. We implemented this analysis using STATA/SE 13.1. We summarize these results in **Supplementary Table 5.**

Behavioral Analyses of Response Times. We used a similar DID framework to analyze the influence of cTBS on response times (rt). We ran a general linear model (GLM) for our regression,

$$rt_{t,c,s,i} = \beta_0 + \beta_1 Task_{(t,c,s,i)} + \beta_2 TMS_{(t,c,s,i)} + \boldsymbol{\phi} Task_{(t,c,s,i)} TMS_{(t,c,s,i)} + \sum_{k=4}^{n} \beta_k X_{(t,c,s,i)}^{k} + \varepsilon_{(t,c,s,i)} \quad (7)$$

We also used cluster-robust standard errors at the subject level. Moreover, we also ran variations of the model and summarized the results in **Supplementary Table 5.**

Hierarchical Bayesian DDM. We analyzed the effect of cTBS on perceptual and value-based decisions using a hierarchical drift diffusion model (HDDM). This decision-making model follows a one-dimensional Wiener process, a dynamical system where the state of evidence, x_t at time t evolves through a stochastic differential equation,

$$\frac{dx_t}{dt} \sim \mathbb{N}(\delta, \sigma^2) \quad (8)$$

where δ is the accumulated evidence at time t,

$$\delta = \kappa_{c,s} E_{c,s,i} \tag{9}$$

where E represents the evidence level and κ is the drift rate that linearly scales the evidence and is typically interpreted as quality of information processing. For the initial conditions (Equation (10)), β represents the start point of the process and we assume that the process takes a decision ρ at time t_d as soon as the evolving variable is either $x_t > \alpha$ (a correct decision) or $x_t \leq 0$ (an incorrect decision). We also accounted for visual sensory processing and motor response delays with the non-decision time parameter, τ. We define reaction time as the sum of non-decision time and the decision time, t_d, $RT = t_d + \tau$.

Given these latent parameters, our goal is to find the Wiener distribution, $\omega(\delta,\alpha,\tau,\beta)$, that best explains the distribution of the empirical choices, $y(\rho,rt)$. To this end, we implemented a hierarchical Bayesian model where each individual point $y_{c,s,i}(\rho,rt)$ follows a Wiener distribution, ω,

$$y_{(c,s,i)} \sim \omega(\delta,\alpha,\tau,\beta) \tag{10}$$

with indices c for task conditions ($c = p$ for perceptual, $c = v$ for value-based), s for participants ($s = 1,\ldots,N_{subjects}$), and i for trials ($i = 1,\ldots,N_{trials}$).

The hierarchical structure contains three random variations at the trial, subject, and condition levels. We treated all interindividual differences per stimulation condition level as random effects:

$$\delta_{(c,s,i)} \sim N\left(\mu_{\delta(s)}E_{(c,s,i)}\ ,\sigma^2_{\delta(s)}\right) \tag{11}$$

$$\tau_{(c,s,i)} \sim N\left(\mu_{\tau(s)},\sigma^2_{\tau(s)}\right) \tag{12}$$

$$\alpha_{(c,s,i)} \sim N\left(\mu_{\alpha(s)},\sigma^2_{\alpha(s)}\right) \tag{13}$$

where $N(\mu, \sigma)$ is a normal distribution with mean, μ, and standard deviation, σ; E are trial-by-trial evidence levels. Using absolute value difference as evidence level, we assume an unbiased fixed starting point of $\beta_{c,s,i} =$ 0.5. Furthermore, we used Bayesian hypothesis testing to compare posterior probability densities.

In summary, we fit a DDM to observed trialwise choices and reaction times to decompose the decision process into distinct latent parameters. These correspond to distinct features of the choice process, which include: (1) the drift rate (δ) or the efficiency of sensory evidence accumulation; (2) the decision boundary (α) or the criteria set to accumulate an amount of evidence to execute a decision; (3) the non-decision time (τ) or the time to process sensory information and execute a motor response; and, (4) the starting point (β) or the bias in the choice process.

Theoretical Accumulated Evidence. We computed estimates for decision times ($t_{d(c,s)}$) and accumulated evidence ($aE_{c,s}$) to test whether aE is a plausible representation of the accumulation process at the neural level. Following the literature (Bogacz et al., 2006, 2010), we define mean decision time as the ratio between the decision threshold and the drift rate shaped by a hyperbolic tangent function,

$$t_{d(c,s)} = \left(\frac{\alpha_{c,s}}{\kappa_{c,s}}\right)\tanh(\kappa_{c,s} \times \alpha_{c,s}) \quad (14)$$

Recall that reaction time, rt, is the sum of both decision and non-decision times, $rt = t_d + \tau$.

Finally, we define accumulated evidence (aE) as the area below the drift process up until the accumulator reaches the decision boundary. We can derive aE using a triangle's area equation, where decision time $t_{d(c,s)}$ is the base and the decision boundary, $\alpha_{c,s}$, is the height,

$$aE_{c,s} = \frac{\alpha_{c,s} \times t_{d(c,s)}}{2} \quad (15)$$

To estimate all parameters, we performed Gibbs sampling via Markov Chain Montecarlo (MCMC) in JAGS (Plummer, 2003, 2016) to generate parameter posterior inferences. We drew a total of 10,000 samples from an initial burn-in step and subsequently drew a total of new 10,000 samples with three chains each. We derived each chain based on different random number generator engines with different seeds. We applied a thinning of 10 to this final sample, resulting in a final set of 1,000 samples for each parameter. This thinning assured auto-decorrelation for all latent variables of interest. We conducted Gelman-Rubin tests for each parameter to confirm chain convergence. All latent parameters in our Bayesian model had $\hat{R} < 1.05$, suggesting that all three chains converged to a target posterior distribution. We compared the difference in posterior population distributions estimated for each parameter between the stimulation conditions (pre-cTBS/post-cTBS) and the differences of these stimulation differences between task conditions (DID, perceptual/value-based). We tested whether the resulting distribution (i.e., the causal stimulation effect) is significantly different from zero (i.e., the null hypothesis) using the cumulative function up to or from 0 depending on the direction of the effect. We refer to this probability as p_{MCMC}.

Hierarchical Bayesian Neural DDM. We also analyzed whether the inclusion of raw trial-by-trial left SFS neural activity in the DDM can improve model evidence. Such a result would suggest that neural activity in the left SFS directly related to the model's latent decision-relevant parameters (Cavanagh et al., 2011; Herz et al., 2016, 2017). We used *z*-scored single-trial neural beta estimates extracted from the left SFS target site. We implemented four a-priori models regarding the role of the left SFS on the decision parameters. Model 1 assumes that the left SFS modulated the decision threshold (**Supplementary Fig. 8a**), while Model 2 assumes that left SFS modulated the drift rate (**Supplementary Fig. 8b**):

$$\alpha_{c,s,i}^{NEURAL} = \alpha_{c,s,i} + \gamma\theta_{c,s,i} \tag{16}$$

$$\delta_{c,s,i}^{NEURAL} = \kappa E_{c,s,i} + \gamma\theta_{c,s,i} \tag{17}$$

where γ is the scale parameter for trial-by-trial left SFS activity, θ. On the other hand, Models 3 and 4 assume that the left SFS modulates both boundary and drift. More specifically, Model 3 assumes separate scale parameters for each latent process (see **Supplementary Fig. 8c**) while Model 4 assumes a common scale parameter for both boundary and drift (see **Supplementary Fig. 8d**). We compared our models' deviance information criterion (DIC) relative to the model without any neural data. Here, the smaller the DIC, the better model performance. We also used the best model for Bayesian post-hoc inferences. Please note that DIC accounts for model complexity (Herz et al., 2016, 2017).

fMRI Data Analysis. Participants performed eight choice-task sessions while BOLD images were recorded with a Philips Achieva 3T whole-body scanner. We used statistical parametric mapping (SPM8, Wellcome Trust Center for Neuroimaging) for image pre-processing and analysis. In particular, images were slice-time corrected (to the acquisition time of the middle

slice) and realigned to account for subject's head motion. Each participant's T1-weighted structural image was co-registered with the mean functional image and normalized to the standard T1 MNI template using the new-segment procedure in SPM8. The functional images were normalized to the standard MNI template using the same transformation, spatially resampled to 3mm isotropic voxels, and smoothed using a Gaussian kernel (FWHM, 8mm).

We estimated two general linear models (GLM), constructed by convolving series of appropriately placed indicator functions with the default model of the BOLD response embedded in SPM8. GLM1 contained only the two indicator functions for the onsets of perceptual or value-based trials. GLM2 contained four indicator functions for the onsets of task (perceptual and value-based trials) and stimulation (pre- or post-stimulation) runs, coupled with one regressor each for parametric modulation of the BOLD response by the trialwise accumulated evidence (aE). Earlier, we have shown that the *theoretical average accumulated evidence* can be derived from subject-level latent DDM parameters by dividing the estimated decision boundary by the estimated drift rate. To construct a trialwise measure of aE, we exploit the fact that the length of the RTs is directly proportional to the size of the decision boundary while the evidence level, E, is directly proportional to the drift rate (Basten et al., 2010; Domenech et al., 2017; Kiani et al., 2014). Given this mapping from subject-wise parameters to trialwise behavioral measures, we can then derive a parametric trialwise measure of accumulated evidence, $aE_{t,c,s,i}$,

$$aE_{t,c,s,i} = \sqrt{\frac{RT_{t,c,s,i}}{E_{t,c,s,i}}} \tag{18}$$

We included a square root function to account for the particular concave nonlinearity in accumulated evidence. Previous work (Tajima et al., 2016) has shown theoretically that the shape of the accumulated evidence is indeed concave, where it suggests that the rate of

accumulating evidence is decreasing as the decision process continues to accumulate. This concavity in the shape of the aE consistent with the predictions from the DDM where evidence strength is steep during earlier responses and flat at later response (Ratcliff and McKoon, 2008; Ratcliff and Smith, 2004).

We convolved our GLMs with a canonical hemodynamic response function, modeled MR image autocorrelations with first-order autoregressive model, and included 6 motion parameters (obtained during realignment) as regressors of no interest. After fitting the model to the BOLD data, we tested regressors for statistical significance at the second-level, in random-effects group one-sample t-tests of the corresponding single-subject contrast images. We performed statistical inference at the cluster level, using a whole-brain family-wise-error-corrected (FWE-corrected) statistical threshold of $p < 0.05$ based on a cluster-forming voxel cutoff at $p < 0.005$ (or $T(19) = 2.9$). For hypothesis-guided region of interest (ROI) analysis, particularly our left SFS stimulation site ($x = -24, y = 24, z = 36$) (Heekeren et al., 2004; Philiastides et al., 2011), we corrected for multiple comparisons using small-volume correction (SVC, $p < 0.05$) restricted within a 10 mm sphere around the target coordinates. We extracted neural betas from this spherical SFS ROI for each participant to perform hypothesis testing and correlational analysis.

Functional Connectivity Analysis. To investigate changes in functional connectivity between the left SFS and other regions in the brain due to cTBS, we ran a psychophysiological interaction (PPI) analysis (Friston et al., 1997). Specifically, we investigated whether brain regions related to accumulated evidence exhibited changes in functional coupling with left SFS as a consequence of the cTBS. To this end, we extracted physiological time series in the left SFS seed region, corresponding to the timecourse of the first eigenvariate across all voxels in the region in a principal component analysis (Friston et al., 1993). The psychological regressor corresponded to the difference in accumulated evidence (as described in GLM2)

between perceptual and value-based decisions. We generated a psycho-physiological interaction (PPI) of the psychological regressors and the time series from the left SFS and computed the PPI contrasts of interest for PDM and VDM. Statistical inference on the subject-specific PPI maps was performed in a second-level random-effects analysis across participants to allow for group-level inferences. For each participant, we also extracted PPI neural betas representing the functional coupling between the left SFS and to perform hypothesis testing and correlational analysis.

Correlating Causal Changes between Neural, Latent, and Behavioral Variables. We tested whether correlations between neural, ν, and behavioral, π, measures had changed after theta-burst stimulation. We define the marginal effect, r, as the correlational change in neural measure, ν, given behavioral measure, π:

$$r(\nu_{c,s}|\pi_{c,s}) = \frac{\partial}{\partial \pi_{c,s}} \mathbb{E}(\nu_{c,s}|\pi_{c,s}) \tag{19}$$

We test the marginal effect, r, of the correlational change between our neural and behavioral measures using our DID framework. In particular, we tested the marginal effect of stimulation on the neural-behavioral correlations at the trial and subject levels. At the trial level, we used a logit regression to test whether the marginal effect of trialwise left SFS neural modulation, $Neur$, will affect choices, ρ. Like previous models, we included various regressors of interest, especially the triple interaction ($Neur \times Task \times TMS$), which measures the causal stimulation effect, $\boldsymbol{\phi}$; and, other regressors, X^k. The full model is,

$$\Pr\left(\rho^{DID}_{(t,c,s,i)}\right) = \left(1 + \exp\left[-\left(\beta_0 + \beta_1 Task_{(t,c,s,i)} + \beta_2 TMS_{(t,c,s,i)} + \beta_3 Neur_{(t,c,s,i)} + \beta_4[Task_{(t,c,s,i)} TMS_{(t,c,s,i)}] + \beta_5[\text{M}\right.\right. \tag{20}$$

Because our model contains a triple interaction term, we removed the nonlinearity bias and isolate the true causal effect, we ran another logit regression without ϕ,

$$\Pr\left(\rho^{NODID}_{(t,c,s,i)}\right) = \left(1 + \exp\left[-\left(\beta_0 + \beta_1 Task_{(t,c,s,i)} + \beta_2 TMS_{(t,c,s,i)} + \beta_3 Neur_{(t,c,s,i)} + \beta_4 [Task_{(t,c,s,i)} TMS_{(t,c,s,i)}] - \right.\right. \quad (21)$$

and we took the difference between the two logit models,

$$\Pr\left(\rho^{TRUEDID}_{(t,c,s,i)}\right) = \Pr\left(\rho^{DID}_{(t,c,s,i)}\right) - \Pr\left(\rho^{NODID}_{(t,c,s,i)}\right). \quad (22)$$

We also ran a GLM to test whether the marginal effect of trialwise left SFS neural betas will causally affect RTs,

$$rt_{(t,c,s,i)} = \beta_0 + \beta_1 Task_{(t,c,s,i)} + \beta_2 TMS_{(t,c,s,i)} + \beta_3 Neur_{(t,c,s,i)} + \beta_4 [Task_{(t,c,s,i)} TMS_{(} \quad (23$$
$$[Neur_{(t,c,s,i)} Task_{(t,c,s,i)}] + \beta_6 [Neur_{(t,c,s,i)} TMS_{(t,c,s,i)}] + \boldsymbol{\phi} [Neur_{(t,c,s,i)} Task_{(t,} \quad)$$
$$TMS_{(t,c,s,i)}] + \varepsilon_{(t,c,s,i)}$$

For both choice and RT models, we also used cluster-robust standard errors at the subject level.

At the subject level, we similarly used linear mixed-effects regression models to test whether the marginal effect of subject-level neural betas (left SFS betas or PPI betas) ν, will affect the behavioral outcomes or DDM-latent parameters, π. Similarly, we estimate the marginal effect, $\boldsymbol{\phi}$, from the three-way interaction, ($\nu \times Task \times TMS$),

$$\pi_{(c,s)} = \beta_0 + \beta_1 Task_{(c,s)} + \beta_2 TMS_{(c,s)} + \beta_3 \nu_{(c,s)} + \beta_4 [Task_{(c,s)} \times TMS] + \beta_5 [\nu_{(c,s)} \times Task_{(c,s)}] + \beta_6 [\nu_{(c,s)} \times TMS_{(c,s)}] + \boldsymbol{\phi} [\nu_{(c,s)} \times Task_{(c,s)} \times TMS_{(c,} + \varepsilon_{(c,s)} \quad (24)$$

In general, the three-way interaction reflects whether the correlations between neural (subject-wise left SFS or PPI betas) and behavioral (choice or DDM parameters) measures are causally affected by stimulation, TMS, and whether the effect is specific only during task, $Task$.

ADDITIONAL REFERENCES:

Ai, C., and Norton, E.C. (2003). Interaction terms in logit and probit models. Economics Letters *80*, 123–129.

Angrist, J.D., and Pischke, J.-S. (2009). Mostly Harmless Econometrics: An Empiricist's Companion (Princeton, NJ: Princeton University Press).

Basten, U., Biele, G., Heekeren, H.R., and Fiebach, C.J. (2010). How the brain integrates costs and benefits during decision making. Proceedings of the National Academy of Sciences *107*, 21767–21772.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How Much Should We Trust Differences-In-Differences Estimates? The Quarterly Journal of Economics *119*, 249–275.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J.D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. Psychological Review *113*, 700–765.

Bogacz, R., Hu, P.T., Holmes, P.J., and Cohen, J.D. (2010). Do humans produce the speed–accuracy trade-off that maximizes reward rate? The Quarterly Journal of Experimental Psychology *63*, 863–891.

Cavanagh, J.F., Wiecki, T. v, Cohen, M.X., Figueroa, C.M., Samanta, J., Sherman, S.J., and Frank, M.J. (2011). Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. Nature Neuroscience *14*, 1462–1467.

Domenech, P., Redouté, J., Koechlin, E., and Dreher, J.-C. (2017). The Neuro-Computational Architecture of Value-based Selection in the Human Brain. Cerebral Cortex 1–17.

Friston, K.J., Frith, C.D., Liddle, P.F., and Frackowiak, R.S.J. (1993). Functional Connectivity: The Principal-Component Analysis of Large (PET) Data Sets. Journal of Cerebral Blood Flow & Metabolism *13*, 5–14.

Friston, K.J., Buechel, C., Fink, G.R., Morris, J., Rolls, E., and Dolan, R.J. (1997). Psychophysiological and Modulatory Interactions in Neuroimaging. NeuroImage *6*, 218–229.

Heekeren, H.R., Marrett, S., Bandettini, P.A., and Ungerleider, L.G. (2004). A general mechanism for perceptual decision-making in the human brain. Nature *431*, 859–862.

Heekeren, H.R., Marrett, S., Ruff, D.A., Bandettini, P.A., and Ungerleider, L.G. (2006). Involvement of human left dorsolateral prefrontal cortex in perceptual decision making is independent of response modality. Proceedings of the National Academy of Sciences *103*, 10023–10028.

Herz, D.M., Zavala, B.A., Bogacz, R., and Brown, P. (2016). Neural Correlates of Decision Thresholds in the Human Subthalamic Nucleus. Current Biology *26*, 916–920.

Herz, D.M., Tan, H., Brittain, J.S., Fischer, P., Cheeran, B., Green, A.L., Fitzgerald, J., Aziz, T.Z., Ashkan, K., Little, S., et al. (2017). Distinct mechanisms mediate speed-accuracy adjustments in cortico-subthalamic networks. ELife *6*, 1–25.

Huang, Y.Z., Edwards, M.J., Rounis, E., Bhatia, K.P., and Rothwell, J.C. (2005). Theta burst stimulation of the human motor cortex. Neuron *45*, 201–206.

Karaca-Mandic, P., Norton, E.C., and Dowd, B. (2012). Interaction terms in nonlinear models. Health Services Research *47*, 255–274.

Kiani, R., Corthell, L., and Shadlen, M.N. (2014). Choice certainty is informed by both evidence and decision time. Neuron *84*, 1329–1342.

Knecht, S., Ellger, T., Breitenstein, C., Ringelstein, E.B., and Henningsen, H. (2003). Changing cortical excitability with low-frequency transcranial magnetic stimulation can induce sustained disruption of tactile perception. Biological Psychiatry *53*, 175–179.

di Lazzaro, V., Pilato, F., Saturno, E., Oliviero, A., Dileone, M., Mazzone, P., Insola, A., Tonali, P.A., Ranieri, F., Huang, Y.Z., et al. (2005). Theta-burst repetitive transcranial magnetic stimulation suppresses specific excitatory circuits in the human motor cortex. The Journal of Physiology *565*, 945–950.

di Lazzaro, V., Ziemann, U., and Lemon, R.N. (2008). State of the art: Physiology of transcranial motor cortex stimulation. Brain Stimulation *1*, 345–362.

Philiastides, M.G., Auksztulewicz, R., Heekeren, H.R., and Blankenburg, F. (2011). Causal role of dorsolateral prefrontal cortex in human perceptual decision making. Current Biology *21*, 980–983.

Puhani, P.A. (2012). The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models. Economics Letters *115*, 85–87.

Plummer, M. (2003) JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Proceedings of the 3rd international workshop on distributed statistical computing 124(125.10): 1-10.

Plummer, M. (2016) rjags: Bayesian graphical models using MCMC. R package version, 4(6).

Ratcliff, R. and Smith, P. (2004) A comparison of sequential sampling models for two-choice reaction time. Psychological Review 111(2): 333-367.

Ratcliff, R. and McKoon, G. (2004) The diffusion decision model: theory and data for two choice-decision tasks. Neural Computation 20(4): 873-922.

Tajima, S., Drugowitsch, J., and Pouget, A. (2016) Optimal policy for value-based decision-making. Nature Communications 7(12400).

Thut, G., and Pascual-Leone, A. (2010). A review of combined TMS-EEG studies to characterize lasting effects of repetitive TMS and assess their usefulness in cognitive and clinical neuroscience. Brain Topography *22*, 219–232.