Transcranial Direct Current Stimulation over the Prefrontal Cortex Alters Encoding and Judgments of Learning Based on Fluency

Alexandra M. Gaynor[1,2] and Elizabeth F. Chua[1,2]

Abstract

■ Past research has shown that judgments of learning (JOLs), subjective confidence judgments made at study about later memorability, are inferential in nature and based on cues available during encoding. Participants tend to use fluency as a cue and give higher JOLs to more fluently encoded items, despite having better recognition memory for disfluently encoded items, which leads to poor JOL accuracy. Research has implicated the dorsolateral prefrontal cortex (DLPFC) and anterior prefrontal cortex (aPFC) in JOL and encoding processes, but no studies to date have tested how the roles of these regions vary with the information on which JOLs are based. We used high-definition transcranial direct current stimulation to test the causal roles of DLPFC and aPFC in encoding success, JOL ratings, and JOL accuracy. Participants studied and made JOLs about words that varied in fluency (i.e., frequency and orientation). High-definition transcranial direct current stimulation over the DLPFC impaired encoding, as evidenced by an increase in subsequent false alarms. For words that were less fluently encoded, aPFC stimulation improved JOL accuracy, perhaps making participants more aware of encoding failures under conditions of disfluency. Conversely, DLPFC and aPFC stimulation decreased JOL accuracy for high-frequency words, suggesting the roles of these regions in JOLs vary with the cognitive bases of the judgments. These results contribute to our understanding of the causal roles of prefrontal regions in objective and subjective memory processes and how their contributions to metamemory accuracy vary with information on which subjective assessments are based. ■

INTRODUCTION

The ability to produce accurate judgments about the status of one's own memory, known as *metamemory monitoring*, is crucial to effective learning and decision-making (Metcalfe, 2002). One of the most commonly studied metamemory monitoring tasks, judgments of learning (JOLs) require individuals to predict, during encoding, whether or not studied stimuli will be successfully remembered at later test (Nelson & Dunlosky, 1991; Nelson & Narens, 1990). How well the predictive judgments correlate with actual subsequent memory performance is known as JOL accuracy. A significant body of research has been conducted on the cognitive bases of JOLs, and it is generally accepted that people make inferences about the memorability of given stimuli based on various cues (Benjamin, Bjork, & Schwartz, 1998; Koriat, 1997). Typically, individuals base JOLs on cues related to encoding fluency, such as speed of encoding (Hertzog, Dunlosky, Robinson, & Kidder, 2003; Benjamin et al., 1998; Begg, Duft, Lalonde, Melnick, & Sanvito, 1989) or word frequency (Benjamin, 2003), despite evidence that higher encoding fluency tends not to benefit memory performance (Koriat & Ma'ayan, 2005; Koriat, 1997; Begg et al., 1989). Conversely, individuals tend to undervalue cues that are predictive of future memory success (Sungkhasettee, Friedman, & Castel, 2011; Shaughnessy, 1981). For instance, participants gave similar JOLs to words studied upright or inverted 180°, but later recalled more inverted than upright words, suggesting participants discount the memory benefit of "desirable difficulties," such as the increased processing required to encode inverted words, which benefits subsequent memory performance (Sungkhasettee et al., 2011). Therefore, immediate JOLs are often inaccurate (Nelson & Dunlosky, 1991) because individuals tend to base confidence on fluency of encoding, which is not predictive of memory success, and fail to consider the memory benefit of deeper encoding due to disfluency (Sungkhasettee et al., 2011; Benjamin, 2003; Koriat, 1997; Begg et al., 1989; Shaughnessy, 1981).

Despite considerable research testing the cognitive bases of JOLs, relatively less is known about the neural mechanisms underlying JOL processes. A small number of neuroimaging and lesion studies provide converging evidence that the pFC is important for JOLs, with different subregions relating to JOL ratings and JOL accuracy

[1]Brooklyn College of the City University of New York, [2]The Graduate Center of the City University of New York

(Kao, Davis, & Gabrieli, 2005; Vilkki, Surma-aho, & Servo, 1999). Lesion studies have broadly implicated pFC in JOLs by demonstrating that, compared with posterior lesion and control participants, patients with pFC lesions make less accurate predictions during study about later retrieval success (Vilkki et al., 1999). Neuroimaging experiments have further identified regions within pFC that may be responsible for specific aspects of JOLs: For JOL ratings, greater activity in the medial pFC and OFC (BA 10 and BA 11) was correlated with higher JOL ratings in a face–name associative encoding task (Do Lam et al., 2012), whereas there was greater activity in lateral (BA 44/BA 6) and ventromedial pFC (vmPFC; BA 11) regions for "will remember" as compared with "will forget" predictions in a scene encoding task (Kao et al., 2005), regardless of JOL accuracy. Similarly, in a verbal encoding task, Yang et al. (2015) found that "will remember" predictions were correlated with greater activity in the dorsolateral pFC (DLPFC; BA 8) and vmPFC (BA 10), as compared with "will forget" predictions. Taken together, findings suggest that both anterior (e.g., orbitofrontal) and posterior (e.g., DLPFC) regions within pFC may track the magnitude of the JOL ratings, and anterior regions associated with JOLs are more medial whereas posterior regions are more lateral. In terms of JOL accuracy, individual differences in JOL accuracy, as indexed by Gamma coefficients, were correlated with vmPFC (BA 11) activity (Kao et al., 2005). Although there is some variation in the subregions whose activity correlates with JOLs due to differences in stimulus type and paradigm, both anterior and posterior pFC subregions have been shown to play roles in JOL ratings and/or accuracy and are ROIs that are examined in the current experiment.

To date, one brain stimulation study has directly compared the roles of anterior and posterior pFC regions in JOLs (Ryals, Rogers, Gross, Polnaszek, & Voss, 2016). Continuous theta burst stimulation (TBS) to the anterior pFC (aPFC) improved JOL accuracy in an associative recognition task, as compared with DLPFC and vertex stimulation, suggesting a causal role of the aPFC in making accurate judgments (Ryals et al., 2016). Although the main effect of stimulation location on trial-by-trial measures of JOL accuracy was marginal, aPFC stimulation decreased overall JOL ratings for subsequent misses, suggesting stimulation made participants more aware of failures during encoding that would result in poor memory performance at test. This raises the hypothesis that perhaps the aPFC supports JOL accuracy by increasing sensitivity to information at encoding that is predictive of later memory performance (Ryals et al., 2016). The current experiment manipulates the cognitive basis of JOL accuracy by testing how JOLs vary for words of different fluencies based on familiarity with the stimulus (i.e., high- vs. low-frequency words) and fluency of perceptual processing (i.e., words presented in an upright or inverted orientation) and investigates how the brain supports JOL accuracy when judgments are made on cues that vary in their influence on memory performance.

Past research has implicated regions of pFC in both memory and metamemory processes: In particular, the DLPFC, which has been shown to play a role in JOLs, is also crucial to successful episodic encoding (Blumenfeld & Ranganath, 2006, 2007; Duarte, Ranganath, & Knight, 2005; Sperling et al., 2001; Wheeler, Stuss, & Tulving, 1995). Given that common brain regions have been associated with both objective and subjective memory processes, it is important to distinguish between the neural underpinnings of JOLs as compared with successful encoding. Some metamemory studies have shown a dissociation between regions that contribute to JOLs and encoding: For instance, patients with pFC lesions show impaired JOLs but intact memory (Vilkki et al., 1999), and fMRI activity in the medial pFC and aPFC has been associated with the process of making JOLs, even when masked with encoding-related activity (Do Lam et al., 2012), suggesting the neural mechanisms underlying JOLs and episodic encoding are at least partially dissociable. Understanding the shared and distinct roles of brain regions in subjective and objective memory is crucial to clarifying the neural bases of JOL accuracy, which reflects the relationship between these two processes. Given that, the aim of the current study is to test the role of pFC in JOLs based on cues that vary in their diagnosticity of memory performance, with a specific focus on how the aPFC and DLPFC may differentially contribute to encoding and JOL processes when the cognitive basis of the JOL varies.

To test the roles of the aPFC and DLPFC in JOLs, we used transcranial direct current stimulation (tDCS), a noninvasive brain stimulation technique, to manipulate brain activity. Conventional tDCS involves passing a weak electrical current from one stimulating electrode, commonly referred to as the "anode," to one return electrode, commonly referred to as the "cathode"; because these electrodes are relatively large and spaced relatively far apart, inferences about specific brain regions being targeted are limited by poor spatial focality, resulting in possible stimulation of regions outside the area of interest (Hampstead, Brown, & Hartley, 2014; Datta, Truong, Minhas, Parra, & Bikson, 2012; Keeser et al., 2011). To improve the spatial focality of tDCS, the current experiment uses "high-definition" tDCS (HD-tDCS), in which smaller electrodes are positioned with one stimulating electrode (i.e., anode) among an array of four return electrodes (i.e., cathodes). Although fewer studies have been conducted using HD-tDCS due to its relative novelty, there is evidence it produces changes in more focal regions of the cortex (Kuo et al., 2013), with stimulation constricted to the area within the radius of the cathodes (Villamar et al., 2013). This increased spatial specificity reduces the likelihood of directly stimulating outside targeted ROIs and makes HD-tDCS a better method for testing how more precise subregions of pFC contribute

to JOL processes. Furthermore, HD-tDCS has been useful at identifying the causal role of the DLPFC in another metamemory task using a semantic retrieval paradigm (Chua, Ahmed, & Garcia, 2017; Chua & Ahmed, 2016). The current experiment uses HD-tDCS to test the causal roles of the aPFC and DLPFC in JOL processes. Specifically, we tested the hypothesis that the aPFC biases individuals to make JOLs based on information that is predictive of later memory success and discount cues that are nondiagnostic of future performance (Ryals et al., 2016). Gaining knowledge about the causal role of the aPFC in JOL accuracy and whether its role varies with the cognitive basis of the judgment is a crucial step toward understanding the neural underpinnings of accurate memory awareness, establishing precise structure–function relationships, and informing future interventions to treat metamemory impairments.

METHODS

Participants

Twenty-five healthy Brooklyn College students consented to participate in this six-session study for financial compensation ($15/hr for 4.5 hr). One participant only completed one session and was withdrawn from the study due to poor impedance at the second session. Thus, data from 24 participants (12 women, age = 18–26 years, M = 21.0, SD = 2.11 years) were included. G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007) was used to determine that, for a repeated-measures ANOVA with one group and three measurements, a sample size of 24 participants was needed for 80% power and a moderate effect size of 0.27, which is lower than the effect size reported in a similar tDCS experiment testing the roles of DLPFC and anterior temporal lobe in metamemory (Chua & Ahmed, 2016). All participants were right-handed, learned English before the age of 5 years, and were free from any self-reported neurological or psychological disorders; medical or skin conditions; unhealed wounds on the scalp, neck, face, or forehead; and metallic implants. All participants gave written consent in a manner approved by the Human Research Protection Program at the City University of New York.

HD-tDCS Protocol

HD-tDCS was delivered using the Soterix 1 × 1 tDCS device (Model 1224-B, Soterix Medical) connected to the Soterix 4 × 1 adapter (Model 4X1-C3 and Model 4X1-C3A, Soterix Medical). All stimulation conditions used five sintered Ag/AgCl ring electrodes (12 mm outer radius and 6 mm inner radius) to deliver low current stimulation. In a within-subjects design, each participant received three stimulation sessions: active stimulation over the left DLPFC, active stimulation over the aPFC, and sham stimulation over the left parietal cortex. Electrode configurations were determined using Soterix HD-Explore (Soterix Medical), with electrode locations corresponding to locations on the 10–20 EEG system. For aPFC stimulation, the "anode" was placed over FPZ with "cathodes" at FP1, FP2, AF3, and AF4 (Figure 1A). For left DLPFC stimulation, the "anode" was placed over F3 with "cathodes" at AF3, F1, F5, and FC3 (Figure 1B). For sham, the "anode" was placed over CP3 with "cathodes" at C3, CP5, CP1, and P3 (not shown).

Before each stimulation session, participants received a "prestimulation tickle," which involved a 30-sec ramp-up to 1 mA and immediate 30-sec ramp down to baseline, to familiarize participants with the sensation of stimulation at half the current intensity of full stimulation and determine whether they were able to tolerate stimulation. All participants were able to tolerate stimulation.

During each active stimulation session, following a 30-sec ramp up, 2 mA of current was delivered through the stimulating electrode and distributed equally among the four return electrodes (0.5 mA each), for the duration of the study task. During sham stimulation sessions, current ramped up to 2 mA and then down to 0.1 mA over

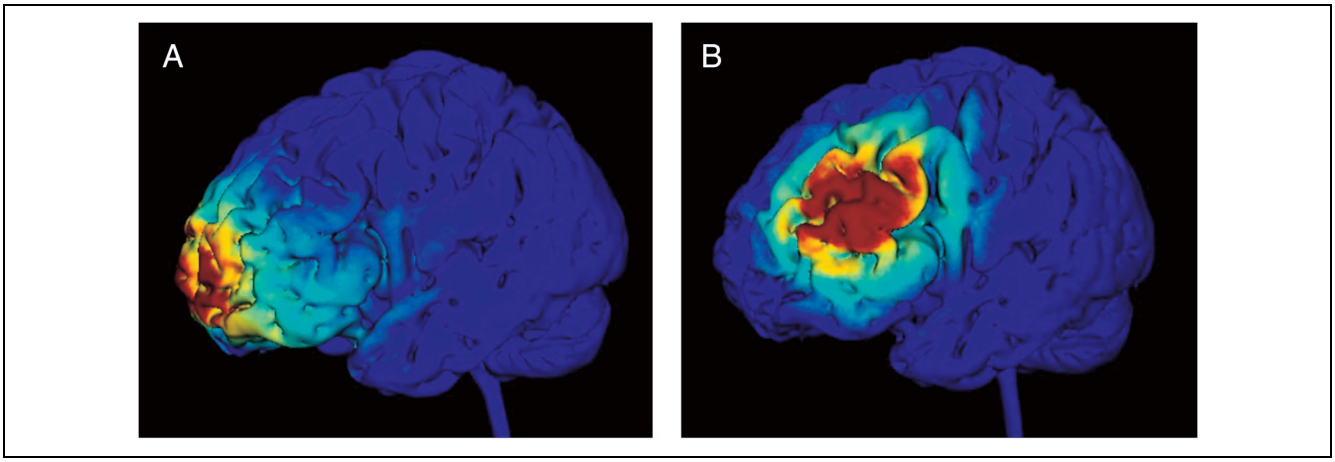

Figure 1. Modeled current densities based on HD-tDCS montages for aPFC (A) and DLPFC (B) using HD-Explore (Soterix Medical).

the course of 30 sec; current was maintained at 0.1 mA for the duration of the task, which is an insufficient strength of current to produce any cognitive effects (Gandiga, Hummel, & Cohen, 2006). The ramp up and down at the start of the sham session mimics the skin sensations experienced by participants during active stimulation (e.g., itching) and has been shown to serve as an effective control in conventional tDCS so participants have difficulty distinguishing between active and sham stimulation (Gandiga et al., 2006, but see Horvath, Carter, & Forte, 2014; Davis, Gold, Pascual-Leone, & Bracewell, 2013; O'Connell et al., 2012). Stimulation began ~3 min before start of the encoding task and was aborted upon completion of the task. Each participant completed six sessions, which consisted of three study–test cycles, with 1 week between each. Study and test were separated by 24 hr to ensure that there were no residual HD-tDCS effects at test. Sessions were counterbalanced for stimulation sites and word lists across participants.

Stimuli and Procedure

Over three study–test sessions, participants were presented with a total of 600 words (average length: 5.63 letters, *SD* = 1.5) from the MRC Psycholinguistics Database (Wilson, 1988), which were divided into three study sets (100 words per set) and three test sets (200 words per set; 100 old/100 new). Fifty percent of the words were presented upright, 50% were presented inverted 180°, and of these, 50% were high-frequency words and 50% were low-frequency words. Word frequency, which refers to the number of times a particular word appears in a given corpus, was determined using the Kucera–Francis Frequency Scale (Kučera & Francis, 1967). Low-frequency words were defined as those with values between 1 and 5, and high-frequency words ranged from 50 to 492, based on past research using Kucera–Francis Frequency ratings to define high- versus low-frequency words (Diana & Reder, 2006; Rudell, 1993). Words were matched across conditions (frequency and orientation) and between lists by word length, concreteness, and imageability.

Study Task

The study task began ~3 min after stimulation began. During the first visit, this time was filled with a practice session, during which participants studied and gave JOLs to six words to familiarize them with the task. During Visits 2 and 3, participants were at rest during the 3 min between the start of stimulation and start of the study task. All stimuli were presented using Psychopy v1.84.0 (Peirce, 2007).

During each study session, participants were presented with 100 words for 2.5 sec each, which were composed of 25 words per frequency/orientation condition: upright/high-frequency, inverted/high-frequency, upright/low-frequency, and inverted/low-frequency. Immediately after presentation of each word, participants were presented with a rating scale from 0 to 100% in 10% increments and gave a JOL rating (2.5 sec) to indicate the likelihood they would remember the previously displayed word 24 hr later at test. Participants gave JOLs using the number scale on the keyboard from left to right, such that "~" = 0% confidence, 1 = 10% confidence, 2 = 20%, 3 = 30%, …, with 0 (i.e., where "10" would be) indicating 100% confidence. Study task duration was 8.3 min, excluding two optional 1-min breaks and self-paced instructions.

After each study session, participants were given a post-stimulation questionnaire consisting of the possible side effects they may have experienced (i.e., headache, neck pain, scalp pain, tingling, burning sensation, skin redness, sleepiness, trouble concentrating, acute mood changes, other), based on guidelines for reporting tDCS effects proposed in previous literature (Brunoni et al., 2011). Participants indicated whether they experienced any of the listed side effects on a scale from 1 to 4 (1 = *absent*, 2 = *mild*, 3 = *moderate*, 4 = *severe*) and whether they believed there was a relationship between the side effect and the stimulation on a scale from 1–5 (1 = *none*, 2 = *remote*, 3 = *possible*, 4 = *probably*, 5 = *definite*). Finally, they were asked to indicate whether they thought they received active or sham stimulation.

Recognition Task

Twenty-four hours after each study session, participants completed a self-paced old/new recognition test. Each test consisted of all 100 studied words, in the same orientation as presented at study, and 100 new words matched on frequency and orientation (25 per frequency/orientation condition). Participants indicated their response via keypress (1 = old, 2 = new). Participants were compensated and debriefed after the final test on their last session.

Data Analyses

To assess memory, we examined the effects of cue type and stimulation on hits and false alarms separately. Past research has shown that, in addition to depth of encoding, greater memory for low- than high-frequency words may also be partly driven by false alarms to high-frequency words at test (Benjamin, 2003). Therefore, it was important to assess false alarms separately rather than in a composite measure of memory such as corrected recognition or d'. Furthermore, we predicted that memory for inverted words would be better than for upright words due to depth of encoding, but perceptual fluency may also cause individuals to produce more false-alarms to upright words at test (Johnston, Dark, & Jacoby, 1985).

JOL accuracy was assessed using d_a, a signal detection theory-based trial-by-trial measure of metacognitive accuracy, as recommended by Benjamin and Diaz (2008) and Masson and Rotello (2009) and used by Chua and

Table 1. Mean Goodman–Kruskal Gamma Coefficients and AUROC2 Measures of JOL Accuracy for Each Condition

	Goodman-Kruskal Gamma Coefficients				*Area Under the Type 2 ROC Curve*			
Stimulation	*High Frequency*	*Low Frequency*	*Upright*	*Inverted*	*High Frequency*	*Low Frequency*	*Upright*	*Inverted*
Sham	0.15 (0.03)	0.12 (0.05)	0.09 (0.04)	0.15 (0.04)	0.56 (0.02)	0.55 (0.02)	0.54 (0.02)	0.56 (0.02)
DLPFC	0.10 (0.03)	0.14 (0.04)	0.14 (0.04)	0.10 (0.04)	0.53 (0.02)	0.55 (0.02)	0.56 (0.02)	0.54 (0.02)
aPFC	0.12 (0.03)	0.25 (0.05)	0.12 (0.04)	0.23 (0.05)	0.55 (0.02)	0.59 (0.01)	0.54 (0.02)	0.60 (0.02)

Standard errors of means in parentheses.

Ahmed (2016), Chua et al. (2017), and Toth, Daniels, and Solinger (2011). There has been a long history of using the Goodman–Kruskal Gamma coefficient to assess metamemory accuracy (Nelson & Narens, 1990), but the measure has been criticized as being suboptimal because it treats data ordinally, does not take into account the magnitude of judgments, and is subject to biases (Masson & Rotello, 2009; Macmillan, Rotello, & Miller, 2004), and many researchers now favor a signal detection-based approach (Fleming, 2017; Maniscalco & Lau, 2014; Masson & Rotello, 2009). d_a measures the distance between two distributions, in this case, JOLs for hits and misses, and is akin to d' values but does not assume equal variance and is thus more appropriate for assessing metacognitive accuracy (Masson & Rotello, 2009). Conceptually, d_a values reflect the difference between the probability of having high confidence in a subsequent hit (metacognitive hit) as compared with high confidence in a subsequent miss (metacognitive false alarm); therefore, a d_a value of 0 would indicate the participant performs at chance, and higher d_a values reflect better metacognitive sensitivity. Because calculating signal detection-based values is problematic when hit or false alarm rates are 0 or 1, we corrected for this using a common approach in which we replaced values of 0 with $0.5/n$ (n = number of trials) and values of 1 with $(n - 0.5)/n$ (Stanislaw & Todorov, 1999). d_a is not the only signal detection theory measure that has been used for metacognitive accuracy; some research reports the area under the Type 2 ROC curve (AUROC2; Baird, Cieslak, Smallwood, Grafton, & Schooler, 2015; Baird, Smallwood, Gorgolewski, & Margulies, 2013; Higham, Perfect, & Bruno, 2009), which is thought to yield similar results as d_a, and other measures, such as meta d', which incorporate both Type 1 and Type 2 d', but are not appropriate for JOLs because JOLs are not given for unstudied items. For ease of future comparison across studies, we have included mean Gamma coefficients, as well as values for AUROC2, in Table 1. The relationships between recognition performance, JOL ratings, JOL accuracy, HD-tDCS site, word frequency, and word orientation were analyzed using mixed linear models and post hoc t tests in SPSS. Full factorial designs examining effects of frequency and orientation were not evaluated due to low trial counts for combined frequency, orientation, and performance conditions (e.g., some individuals had no high-frequency upright misses; see Table 2). Instead, separate models were used to analyze the effects of (1) frequency and stimulation on recognition performance, JOL ratings, and d_a and (2) the effects of orientation and stimulation on recognition performance, JOL ratings, and d_a.

Previous research has suggested the effects of tDCS may vary based on a number of individual differences, including anatomical differences related to sex, age, and head size (Datta et al., 2012). Furthermore, research has shown that internal psychological states, such as transient changes in mood (Harrison et al., 2008; Mayberg et al., 1999), alertness (Braboszcz & Delorme, 2011), and motivation (Berryhill, Peterson, Jones, & Stephens, 2014) alter baseline neural activity, which is likely to mediate the effects of tDCS (Learmonth, Thut, Benwell, & Harvey, 2015; Berryhill et al., 2014; Neuling, Rach, & Herrmann, 2013). Thus, to control for the effects of individual variation in responses to tDCS, we included variables reflecting participants' sex, age, head size, education level, mood, alertness, and sensitivity to stimulation as covariates in these mixed linear models; all covariates in each model were evaluated at values of 0. Participant ID was included as a random effect. Stimulation site was coded with aPFC as the reference, and all other predictors were mean-centered to allow for interpretation of the intercept and avoid multicollinearity when assessing interactions. To take into account the number of fixed-effects parameters being estimated, all models used a restricted maximum likelihood procedure (SPSS Version 23.0) to yield unbiased parameter estimates. One participant was excluded from all analyses of JOL accuracy based on outlier d_a performance during sham stimulation (d_a = 0.99), and another participant was excluded from analyses of JOL accuracy that included frequency as a predictor, due to extreme outlier performance for low-frequency words (d_a = −3.29). Therefore, following exclusion of these outliers, sample size for analyses of orientation as a predictor of d_a was 23 participants, and sample size for analyses of frequency as a predictor of d_a was 22 participants.

Table 2. Trial Counts for Each Participant for All Conditions in Full Factorial Design (Frequency × Orientation × Recognition Performance × Stimulation Site)

	Sham								*DLPFC*								*aPFC*							
	HF Upright		*HF Inverted*		*LF Upright*		*LF Inverted*		*HF Upright*		*HF Inverted*		*LF Upright*		*LF Inverted*		*HF Upright*		*HF Inverted*		*LF Upright*		*LF Inverted*	
Participant	*Hit*	*Miss*	*Hit*	*Miss*	*Hit*	*Miss*	*Hit*	*Miss*	*Hit*	*Miss*	*Hit*	*Miss*	*Hit*	*Miss*	*Hit*	*Miss*	*Hit*	*Miss*	*Hit*	*Miss*	*Hit*	*Miss*	*Hit*	*Miss*
01	24	1	23	2	24	0	21	4	24	1	24	1	25	0	23	2	25	0	23	2	21	3	20	4
02	20	3	22	3	22	3	21	4	23	2	22	3	19	6	21	4	16	4	21	2	18	5	16	8
03	16	9	20	4	20	5	21	4	17	7	16	7	18	6	21	2	18	6	19	3	14	11	19	4
04	19	6	20	5	18	7	21	4	8	17	14	11	11	13	15	10	22	3	18	7	19	6	21	4
05	18	7	20	5	19	6	20	4	17	8	22	3	16	8	19	6	15	10	19	6	20	5	16	9
06	8	17	7	18	8	17	8	17	14	11	14	11	17	8	9	16	6	19	7	17	5	20	8	17
07	12	12	15	8	18	6	19	2	12	10	9	14	9	13	12	11	8	17	14	11	12	13	12	13
08	14	11	18	6	17	8	13	11	6	19	5	20	11	14	16	8	12	12	17	8	15	10	8	17
09	10	15	14	11	16	8	20	5	20	5	21	4	17	8	21	4	20	5	22	3	24	1	20	5
10	15	10	20	5	19	6	19	5	10	15	20	5	19	6	20	5	17	8	20	5	19	6	22	3
11	19	6	14	9	19	6	22	3	20	4	22	3	15	10	14	11	11	14	7	18	7	18	13	11
12	21	4	23	2	21	4	23	1	24	1	20	4	21	4	21	3	23	2	22	3	18	6	23	1
13	15	9	18	7	17	7	19	4	8	16	17	8	17	7	21	3	13	10	15	9	14	11	21	4
14	11	13	15	9	9	14	14	11	19	6	20	5	19	6	18	6	18	6	17	4	19	5	21	4
15	6	18	6	19	6	17	9	16	9	15	8	16	8	17	12	13	3	22	7	15	8	14	11	13
16	13	12	12	13	14	11	17	8	11	14	9	16	10	15	7	17	17	8	21	4	16	9	20	5
17	19	6	18	7	16	9	18	7	12	13	18	5	17	8	16	9	20	5	19	6	11	14	17	8
18	8	17	16	9	13	10	10	10	11	12	21	3	13	10	22	2	10	14	17	8	10	11	17	7
19	20	4	23	1	21	4	23	2	23	1	24	1	20	4	23	1	25	0	23	1	23	2	24	1
20	20	5	24	1	21	4	20	5	17	8	21	4	16	9	24	0	24	1	22	3	22	3	24	1
21	9	16	12	13	14	10	14	11	9	16	18	7	16	9	16	9	14	11	12	13	10	15	9	16
22	13	10	12	11	14	11	14	8	13	11	9	15	16	8	18	4	10	15	16	8	16	8	15	5
23	4	20	6	15	6	17	9	12	8	16	11	12	6	17	9	15	13	11	9	11	8	16	7	14
24	13	12	9	16	14	11	16	9	19	6	18	7	15	9	13	10	18	7	16	9	11	13	15	10

Low trial counts under several conditions prevented analysis of full factorial design due to inability to accurately calculate d_a measures of JOL accuracy. HF = high-frequency; LF = low-frequency.

For all other analyses, $n = 24$. All results were considered significant at $p < .05$.

RESULTS

Task Duration

Because participants could take optional breaks during encoding and stimulation was terminated at the end of the encoding, we first evaluated whether there were differences in stimulation duration using a repeated-measures ANOVA. Total study task duration ranged from 9.24 to 11.28 min ($M = 9.62$ min, $SD = 0.55$ min), and stimulation site did not significantly affect time to complete study, $F(2, 46) = 0.340, p = .713$. Time to complete the recognition test ranged from 4.61 to 10.88 min ($M = 7.56$ min; $SD = 1.41$), with no differences in duration between stimulation conditions, $F(2, 46) = 0.183, p = .835$.

Participant Blinding

To assess whether participants were blinded to the stimulation condition (i.e., active vs. sham), we examined their responses on the post-stimulation sensation questionnaire about whether they thought they received active or sham stimulation. In the sham condition, 14 participants correctly guessed they received sham stimulation, and 10 incorrectly guessed they received active stimulation. In the DLPFC condition, 17 participants correctly guessed they received active stimulation, and 7 incorrectly guessed they received sham stimulation. In the aPFC condition, 17 participants correctly guessed they received active stimulation, and 7 incorrectly guessed they received sham stimulation. To test if the belief about stimulation differed between conditions, we used a repeated-measures logistic regression with aPFC stimulation as a reference. Significantly more participants guessed they received active stimulation during aPFC stimulation as compared with sham (Wald $\chi^2 = 5.549, p < .05$) and during DLPFC compared with sham (Wald $\chi^2 = 3.953, p < .05$), but there was no significant difference in participants' guesses that they received active stimulation between DLPFC and aPFC sessions (Wald $\chi^2 = 0.400, p = .527$).

Memory Performance

Subsequent Recognition—Word Frequency

To understand any potential effects of stimulation on recognition, we examined changes in hits and false alarms separately and constructed two models with stimulation and frequency as predictors for each trial type (Figure 2) and included individual differences as covariates. For hits (Figure 2A), word frequency was a significant predictor, $F(1, 251.211) = 18.536, p < .001$, with higher hit rates for low-frequency words ($M = 0.660$, $SE = 0.038$) compared with high-frequency words ($M = 0.618$, $SE = 0.038$; mean difference = 0.042, 95% CI [0.023, 0.061]), as expected due to deeper encoding of low-frequency words. There was no main effect of stimulation site ($p = .371$) and no significant interaction between stimulation site and word frequency ($p = .311$).

Because the recognition benefit of low-frequency words has also shown to be driven by higher false alarms for high-frequency words (Benjamin, 2003), we analyzed the effects of Stimulation and Word frequency on false alarm rates (Figure 2B), again including individual differences as covariates. There was a significant main effect of Stimulation, $F(2, 254.108) = 5.270, p < .01$; a significant main effect of Frequency, $F(1, 251.381) = 57.946, p < .001$; and no significant interaction between Stimulation and Frequency on mean predicted false alarms ($p = .798$). Pairwise comparisons testing the significant main effect of Stimulation site showed that DLPFC stimulation significantly increased subsequent false alarms ($M = 0.267$, $SE = 0.027$) compared with false alarm rates following sham stimulation ($M = 0.241$, $SE = 0.027$; mean difference = 0.026, 95% CI [0.002, 0.049], $p < .05$) and following aPFC stimulation ($M = 0.230$, $SE = 0.027$; mean difference = 0.037, 95% CI [0.014, 0.060], $p < .01$). There was no difference in false alarm rates between aPFC and sham ($p = .386$). There was also a significant main effect of Word frequency: As expected, participants had higher false alarm rates for high-frequency words ($M = 0.279$, $SE = 0.027$) than low-frequency words ($M = 0.213$, $SE = 0.027$; mean difference = 0.067, 95% CI [0.049, 0.084], $p < .01$). Taken together, these results suggest that participants benefit from the distinctiveness of low-frequency words at encoding, as shown by higher hit rates for low- vs. high-frequency words, but encoding was impaired by DLPFC stimulation, as shown by an increase in subsequent false alarm rates.

Subsequent Recognition—Word Orientation

We next constructed a model to examine the effects of stimulation and word orientation as predictors of hits and false alarms. A model including Stimulation and Orientation as predictors of hit rates with individual differences as covariates showed no significant main effect of Stimulation, $F(2, 252.779) = 1.120, p = .328$, and no interaction between Stimulation and Orientation, $F(2, 251.186) = 0.528, p = .590$, but a significant main effect of Orientation on predicted hit rates, $F(1, 251.186) = 43.962, p < .001$. Hit rates for inverted words ($M = 0.669$, $SE = 0.038$) were significantly higher than hit rates for upright words ($M = 0.609$, $SE = 0.038$; mean difference = 0.061, 95% CI [0.043, 0.079]; Figure 3A), suggesting subsequent memory benefitted from greater depth of encoding and distinctiveness of inverted words at encoding.

Turning to false alarm rates, there was a significant main effect of Stimulation site, $F(2, 254.243) = 5.025$, $p < .01$; a significant main effect of Orientation, $F(1, 251.399) = 5.534, p < .05$; and no interaction between

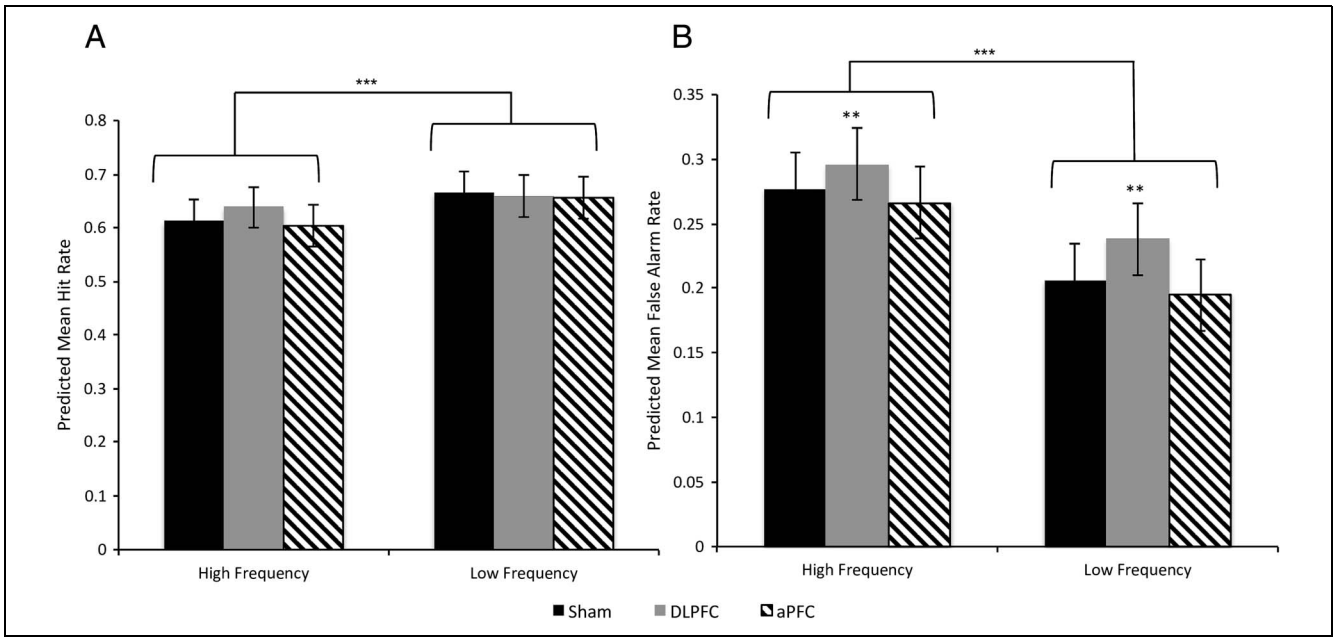

Figure 2. Effects of frequency and stimulation on recognition. Participants had higher hit rates for low-frequency than high-frequency words (A). Participants had higher false alarms for high- than low-frequency words and higher false alarm rates following DLPFC stimulation relative to sham and aPFC stimulation (B). Error bars reflect standard errors of the means. *Y* axes reflect mean hit and false alarm rates predicted by mixed linear model. ***Main effect of Frequency, $p < .001$. **Main effect of Stimulation, $p < .01$ with DLPFC > sham and aPFC ($p < .01$).

Stimulation and Orientation on predicted false alarm rates, $F(2, 251.399) = 0.024$, $p = .976$. False alarms for upright words ($M = 0.257$, $SE = 0.027$) were significantly higher than false alarms for inverted words ($M = 0.236$, $SE = 0.027$; mean difference = 0.021, 95% CI [0.003; 0.039]; Figure 3B). Pairwise comparisons on the effects of stimulation on false alarms showed that false alarms were significantly higher following DLPFC stimulation ($M = 0.267$, $SE = 0.027$) relative to sham ($M = 0.241$, $SE = 0.027$; mean difference = 0.026, 95% CI [0.001, 0.050], $p < .05$) and aPFC stimulation ($M = 0.230$, $SE = 0.027$, mean difference = 0.037, 95% CI [0.013, 0.061], $p < .01$), with no significant difference between sham and aPFC ($p = .395$). Thus, similar to the effects of stimulation on memory

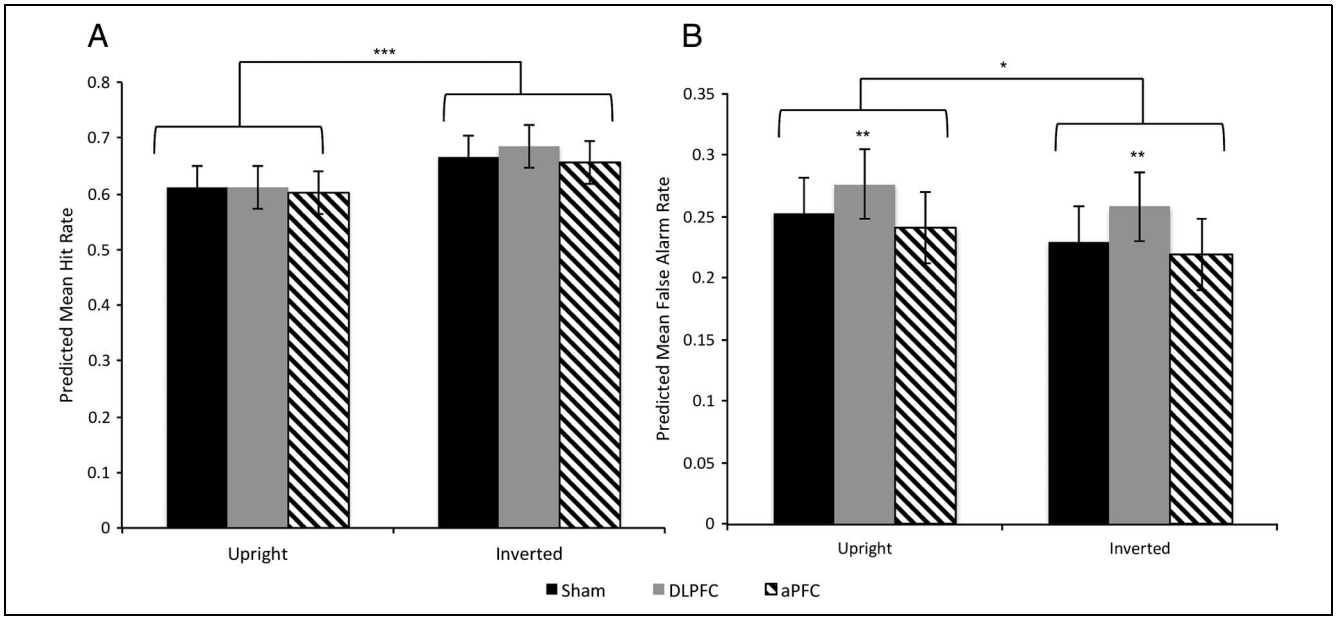

Figure 3. Effects of orientation and stimulation on recognition. Hit rates for inverted words were greater than hit rates for upright words (A). False alarm rates were significantly higher for upright than inverted words and false alarm rates were greater with DLPFC than sham and aPFC stimulation (B). Error bars reflect standard errors of the means. *Y* axes reflect mean hit and false alarm rates predicted by mixed linear model. ***Main effect of orientation on hit rate, $p < .001$. **Main effect of Stimulation on false alarms, $p < .01$ with DLPFC > sham ($p < .05$) and DLPFC > aPFC ($p < .01$). *Main effect of Orientation on false alarms, $p < .05$.

for high- and low-frequency words, DLPFC stimulation increased false alarm rates. Taken together, these results suggest encoding was poorer under DLPFC stimulation, as evidenced by poorer ability to distinguish between "new" and "old" items at test, regardless of frequency/orientation.

Metamemory Performance

Mean JOL Ratings—Word Frequency

A primary aim of the current experiment was to test the effects of stimulation on the magnitude and accuracy of JOLs based on different cue types. Across all sessions, participants' mean JOLs ranged from 38.30 to 88.59 (M = 63.99, SD = 13.33). We first constructed a model using Stimulation site, Word frequency, and Subsequent performance on mean JOL ratings with individual differences as covariates. There was a significant main effect of Subsequent recognition, $F(1, 244.781) = 31.744$, $p < .001$, and a significant main effect of Frequency, $F(1, 244.781) = 34.218$, $p < .001$. There was no significant main effect of Stimulation, $F(2, 247.462) = 0.297$, $p = .744$, and no two- or three-way interactions.

Subsequent recognition performance was a significant predictor of mean JOL ratings: As expected, mean JOLs for subsequently remembered items (M = 65.525, SE = 3.068) were significantly higher than mean JOLs for subsequent misses (M = 59.947, SE = 3.068; mean difference = 5.578, 95% CI [3.628, 7.528]). Word frequency was also a significant predictor of mean JOLs as predicted; JOLs for high-frequency words (M = 65.632, SE = 3.068) were significantly higher than for low-frequency words (M = 59.841, SE = 3.068; mean difference = 5.791, 95% CI [3.841, 7.742]). Results are consistent with past work (Jia et al., 2016; Sungkhasettee et al., 2011; Benjamin, 2003) and support our hypothesis that participants base JOLs on fluency at encoding. Stimulation site did not predict differences in mean JOL ratings, and there were no significant interactions in the effects of stimulation, frequency, and performance on mean JOLs.

Mean JOL Ratings—Word Orientation

We then constructed a model using Stimulation site, Word orientation, and Subsequent performance on mean JOL ratings with individual differences as covariates. Again, as predicted, participants gave higher JOLs to subsequent hits (M = 65.493, SE = 3.116) than subsequent misses (M = 60.350, SE = 3.116; mean difference = 5.143, 95% CI [3.316, 6.970]), $F(1, 244.816) = 30.748$, $p < .001$. Consistent with research showing participants often incorrectly base confidence on fluency of encoding (Koriat, 1997; Johnston et al., 1985), participants also gave higher JOLs to upright words (M = 64.026, SE = 3.116) than inverted words (M = 61.817, SE = 3.116; mean difference = 2.210, 95% CI [0.383, 4.037]), $F(1, 244.816) = 5.675$, $p < .05$. However, Stimulation did not significantly predict mean JOLs, $F(2, 247.121) = 0.414$, $p = .661$, and there were no significant two- or three-way interactions. Taken together, results demonstrate typical findings regarding the cues used to make JOL ratings: Participants gave higher JOLs to high-frequency and upright words, suggesting they base confidence in retrieval on ease of encoding (Koriat, 1997), but Stimulation had no effect on JOLs for any cue type.

JOL Accuracy—Word Frequency

As is typical with immediate JOLs, participants' JOL accuracy was poor to moderate, with mean d_a values across all cue conditions ranging from −0.28 to 0.73 (M = 0.22, SD = 0.23). To understand how stimulation affected JOL accuracy and whether this varied by word frequency, we constructed a mixed linear model using Stimulation site and Frequency as predictors of d_a and included individual differences as covariates. There was a main effect of Word frequency on predicted mean d_a, $F(1, 490.084) = 14.756$, $p < .001$, with JOL accuracy being significantly higher for low-frequency words (M = 0.258, SE = 0.055) than high-frequency words (M = 0.175, SE = 0.055; mean difference = 0.083, 95% CI [0.040, 0.125]). There was also a main effect of Stimulation site, $F(2, 498.879) = 7.642$, $p < .001$, with worse JOL accuracy in the DLPFC condition (M = 0.155, SE = 0.056) compared with sham (M = 0.229, SE = 0.057; mean difference = −0.074, 95% CI [−0.033, −0.016], $p < .05$) and compared with aPFC stimulation (M = 0.265, SE = 0.056; mean difference = −0.110, 95% CI [−0.168, −0.053], $p < .001$), with no difference in d_a between aPFC and sham conditions ($p = .258$). This main effect of stimulation was qualified by a significant interaction between stimulation site and word frequency, $F(2, 490.084) = 11.780$, $p < .001$, showing differential effects of stimulation depending on word frequency. For low-frequency words, aPFC stimulation significantly improved d_a values relative to sham (mean difference = 0.164, 95% CI [0.083, 0.245], $p < .001$) and relative to DLPFC stimulation (mean difference = 0.181, 95% CI [0.104, 0.259], $p < .001$). In contrast, for high-frequency words, aPFC and DLPFC stimulation significantly impaired accuracy relative to sham (aPFC vs. sham: mean difference = −0.092, 95% CI [−0.173, −0.011], $p < .05$]; DLPFC vs. sham: mean difference = −0.132, 95% CI [−0.210, −0.054], $p < .001$; Figure 4A). In summary, these results suggest aPFC stimulation selectively improved JOL accuracy for the less fluent condition, whereas both pFC stimulation sites decreased accuracy for words that were more fluent at encoding.

JOL Accuracy—Word Orientation

To test whether stimulation affected JOL accuracy when judgments were based on word orientation, we constructed a mixed linear model using Stimulation site and Word orientation as predictors of d_a, including

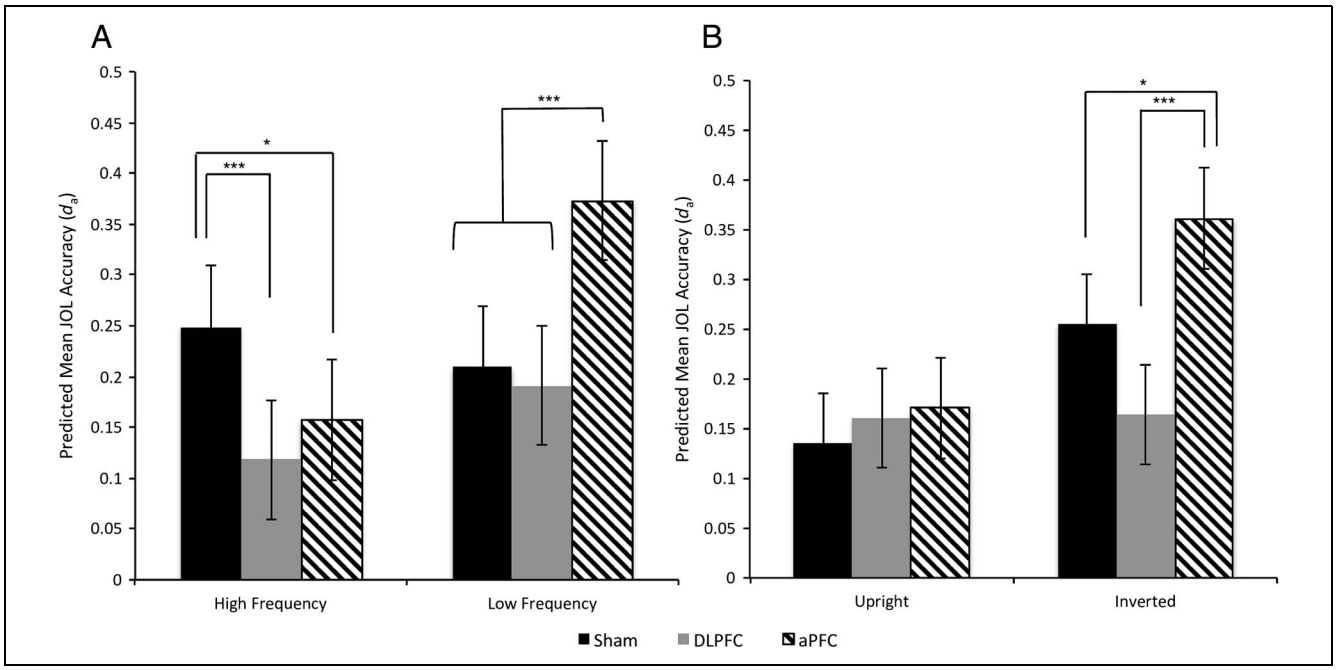

Figure 4. Effects of stimulation and frequency/orientation on JOL accuracy. For low-frequency words, aPFC stimulation improved JOL accuracy relative to sham and DLPFC stimulation, whereas for high-frequency words, aPFC and DLPFC stimulation impaired JOL accuracy for high-frequency words relative to sham (A). For inverted words, aPFC stimulation improved accuracy for inverted words relative to sham and DLPFC stimulation, but there was no effect of Stimulation on upright words (B). Error bars reflect standard errors of the means. *y* axes reflect mean JOL accuracy predicted by mixed linear model. Asterisks denote significant post hoc tests following significant interaction effects: ***$p < .001$. **$p < .01$. *$p < .05$.

individual differences as covariates. There was a significant main effect of Word orientation, $F(1, 513.516) = 13.418$, $p < .001$; a significant main effect of Stimulation site, $F(2, 531.256) = .796$, $p < .05$; and a significant interaction between Orientation and Stimulation site, $F(2, 513.516 = 3.666$, $p < .05$, on predicted mean d_a. Predicted mean d_a was higher for inverted words ($M = 0.260$, $SE = 0.041$) relative to upright words ($M = 0.156$, $SE = 0.041$; mean difference = 0.104, 95% CI [0.048, 0.160]). There was a significant main effect of Stimulation site on JOL accuracy, and pairwise comparisons demonstrated that d_a for aPFC stimulation ($M = 0.266$, $SE = 0.044$) was marginally higher than during sham stimulation ($M = 0.195$, $SE = 0.045$; mean difference = 0.071, 95% CI [0.010, 0.152], $p = .084$), and significantly higher than during DLPFC stimulation ($M = 0.162$, $SE = 0.044$; mean difference = 0.104, 95% CI [0.029, 0.179], $p < .01$). This main effect of Stimulation was qualified by a significant interaction between Stimulation site and Word orientation on mean predicted d_a. Pairwise comparisons revealed that Stimulation site had no effect on d_a values for upright words. For inverted words, relative to sham ($M = 0.255$, $SE = 0.051$), aPFC stimulation significantly improved JOL accuracy ($M = 0.361$, $SE = 0.051$; mean difference = 0.106, 95% CI [0.000, 0.212], $p < .05$) and DLPFC stimulation marginally impaired JOL accuracy ($M = 0.164$, $SE = 0.050$; mean difference = −0.091, 95% CI [−0.194, −0.11], $p = .080$). This produced a significant difference in JOL accuracy for inverted words between DLPFC and aPFC stimulation sites (mean difference = 0.197, 95% CI [0.096, 0.299], $p < .001$; Figure 4B). Similar to the effects of stimulation on d_a for high- and low-frequency words, here, aPFC stimulation appears to have selectively improved JOL accuracy for words in the less fluent condition.

Exploratory Analyses

Although analyses of d_a showed that aPFC stimulation improved accuracy for disfluent conditions, we did not find a significant interaction between the effects of stimulation, fluency, and accuracy on mean JOLs (i.e., there was no difference in the difference between mean JOLs for hits and misses based on fluency or stimulation), which raises the question of what scaling behaviors drove the improvement in JOL accuracy. We reasoned that one possibility was that JOLs were distributed differently during different stimulation conditions and that means may not have been sensitive to this. Therefore, we conducted exploratory analyses testing the differences in median JOLs, rather than means, for low-fluency conditions by stimulation site. There were no differences between stimulation conditions for hits, and the effects on metamemory accuracy appear to be driven by JOLs given to the misses, similar to prior work (Ryals et al., 2016). For low-frequency misses, median JOLs were lower under aPFC than DLPFC stimulation ($p < .05$) and marginally lower under aPFC than sham stimulation ($p = .069$). For inverted misses, results were insignificant but trending in a similar direction, with JOLs lower under aPFC than DLPFC stimulation ($p = .11$) and lower under aPFC than sham stimulation

($p = .086$). Thus, participants may have shifted some of their JOLs to the lower end of the rating scale for low-frequency and inverted misses under aPFC stimulation. To clarify why this was not reflected in the analyses of mean JOLs, we analyzed the distribution of JOLs by calculating skewness for each stimulation and fluency condition and found that nearly all JOL distributions were negatively skewed, suggesting participants used the higher end of the confidence scale, but this was not reflected by the means. However, for low-fluency misses under aPFC stimulation, the distributions were less negatively skewed; therefore, it appears that participants had greater awareness of encoding failures when encoding was perceived as difficult (i.e., disfluent conditions) and appropriately gave more JOL responses on the lower end of the rating scale for subsequent misses.

DISCUSSION

Although the aPFC has been implicated in JOL accuracy (Ryals et al., 2016; Kao et al., 2005) and the DLPFC has been associated with JOL magnitude and encoding success (Do Lam et al., 2012; Blumenfeld & Ranganath, 2007; Kao et al., 2005), no research to date has investigated how the roles of these regions may vary with the cognitive bases of JOLs. We replicated previous behavioral findings by showing that mean JOL ratings were higher for high-frequency and upright words, consistent with an "easily learned, easily remembered" heuristic (Koriat, 1997). However, because memory performance benefits from depth of encoding (Bjork & Bjork, 2011; Craik & Lockhart, 1972) and stimulus novelty (Kishiyama, Yonelinas, & Knight, 2009), recognition was better for low-frequency and inverted words. Turning to the effects of stimulation on JOLs and encoding, HD-tDCS over the DLPFC impaired encoding, as evidence by increased false alarms for all cue types, and stimulation over the aPFC selectively improved JOL accuracy for the least fluent encoding conditions, that is, low-frequency and inverted words. Collectively, our results suggest a causal role of the DLPFC in encoding, indicate that memory and metamemory functions are at least partially dissociable in pFC, and suggest the contributions of the aPFC and DLPFC in metamemory accuracy vary based on encoding fluency.

Behavioral Effects of Cue Type on Memory and Metamemory

Our findings that word frequency and orientation predicted recognition performance are consistent with past research suggesting participants have better memory under conditions of encoding disfluency (Besken & Mulligan, 2013; Yue, Castel, & Bjork, 2013), including disfluency that results from lack of familiarity with the stimulus (Jia et al., 2016; Kishiyama & Yonelinas, 2003; Whittlesea & Williams, 2000). Several studies have shown that the novelty of verbal stimuli influences recognition success (Kishiyama et al., 2009; Kishiyama & Yonelinas, 2003), likely because novelty leads to deeper processing or attentional orienting to unfamiliar stimuli (Kishiyama et al., 2009; Corbetta & Shulman, 2002). In the current study, participants had higher hit rates and lower false alarm rates for low-frequency as compared with high-frequency words, confirming the benefit of relative novelty on recognition performance. We found similar effects of word orientation on memory performance: Participants had higher hit rates and lower false alarm rates for inverted as compared with upright words, consistent with past research suggesting perceptual disfluency promotes memory success due to deeper encoding (Besken & Mulligan, 2013; Rhodes & Castel, 2008).

Despite having better overall memory for low-frequency and inverted words, JOLs were significantly higher for high-frequency and upright words. This is consistent with past research showing the effects of word frequency on metamemory ratings (Jia et al., 2016; Sungkhasettee et al., 2011; Benjamin, 2003) and supports the notion that participants use an "easily learned, easily remembered" heuristic, wherein confidence increases with fluency of encoding (Besken & Mulligan, 2013; Miele, Finn, & Molden, 2011; Koriat, 2008). Because low-frequency words are more disfluent due to lower baseline familiarity with the stimulus (Jia et al., 2016; Balota, Burgess, Cortese, & Adams, 2002) and processing of inverted words is more effortful due to perceptual disfluency (Rhodes & Castel, 2008; Johnston et al., 1985), participants rely on fluency as a cue to learning and mistakenly give lower JOLs to words under these conditions, based on the subjective experience of more effortful encoding. This is consistent with research using a wide variety of paradigms, which show in conditions with less fluent encoding, such as difficult-to-read fonts (Yue et al., 2013; Diemand-Yauman, Oppenheimer, & Vaughan, 2011), elaborative rehearsal (Shaughnessy, 1981), and interactive imagery (Rabinowitz, Ackerman, Craik, & Hinchley, 1982), there is better learning because items are more deeply processed, but individuals tend to discount these memory benefits when judging their encoding success (Yue et al., 2013; Bjork & Bjork, 2011; Sungkhasettee et al., 2011).

Effects of HD-tDCS on Memory and Metamemory

In addition to the behavioral effects of frequency and orientation on recognition performance, we showed that HD-tDCS over the DLPFC led to greater subsequent false alarms, as compared with aPFC and sham stimulation. This pattern held for the model that included frequency, with greater false alarm rates for both high- and low-frequency words, and for the model that included orientation, with greater false alarm rates for both upright and inverted words. These results suggest the DLPFC plays a causal role in encoding success by supporting accurate discrimination between old and new items, and stimulation interfered with this function.

Most past literature implicating the DLPFC in false alarms has focused on its contribution during retrieval (Yonelinas, Otten, Shaw, & Rugg, 2005; Henson, Rugg, Shallice, Josephs, & Dolan, 1999; Parkin, Bindschaedler, Harsent, & Metzler, 1996), but some research has shown the lateral pFC during encoding is associated with subsequent false alarms (Demeter, Mirdamadi, Meehan, & Taylor, 2016; Slotnick & Schacter, 2004). One fMRI study showed that increased activity in the lateral pFC (BA 10/BA 45) during encoding of shapes was associated with subsequent false alarms to similar nonstudied shapes as compared with subsequent hits (Slotnick & Schacter, 2004), although it is worth noting that these regions are more ventral/anterior than our DLPFC stimulation site. However, another study showed that excitatory short TBS to the left DLPFC during encoding marginally lowered the proportion of subsequent false alarms for lure items relative to stimulating the vertex (Demeter et al., 2016), indicating that perhaps short TBS to the DLPFC strengthened encoding of specific information, which led to a decrease in false alarms. Our finding that DLPFC stimulation increased false alarms is consistent with research implicating this region in encoding processes that predict subsequent false alarms.

It is worth noting that our results show that tDCS to the DLPFC during encoding disrupted, rather than enhanced, recognition performance by increasing false alarms. Although some studies have shown that tDCS to pFC can benefit memory performance (Chua et al., 2017; Matzen, Trumbo, Leach, & Leshikar, 2015), other studies have shown that stimulation fails to improve performance in other cognitive domains (Gaynor & Chua, 2017; Boggio et al., 2010; Monti et al., 2008; Marshall, Mölle, Siebner, & Born, 2005), and this is likely due to differences in task design, stimulation parameters such as electrode montages, and characteristics of the participants in the sample (Tremblay et al., 2014). Therefore, our finding that tDCS to the DLPFC impaired subsequent recognition is consistent with research suggesting that excitatory stimulation does not always have facilitatory effects on performance (Gaynor & Chua, 2017; Tremblay et al., 2014; Boggio et al., 2010; Monti et al., 2008; Marshall et al., 2005) and demonstrates a causal role of the DLPFC in encoding success.

Based on past research implicating the aPFC in JOL accuracy (Ryals et al., 2016), we hypothesized that aPFC stimulation would bias individuals toward using cues that are predictive of memory success, leading to better JOL accuracy in the low-frequency and inverted conditions. Indeed, aPFC stimulation significantly improved JOL accuracy, as measured by d_a, for low-frequency and inverted words, relative to sham and DLPFC stimulation. Although aPFC stimulation improved JOL accuracy under disfluent conditions, we also found that both DLPFC and aPFC stimulation decreased JOL accuracy for high-frequency words relative to sham, suggesting the contribution of these regions varies based on the diagnosticity of the cue: For disfluent conditions, the aPFC made participants' JOLs more sensitive to encoding failures, but in the case of high-frequency words, which are familiar and fluently encoded, DLPFC and aPFC stimulation made participants more reliant on the nondiagnostic cue of fluency. The finding that the aPFC increased awareness of memory performance for inverted and low-frequency words is in line with the theory that, when faced with multiple possible responses based on bottom–up information from posterior regions, anterior portions of pFC act to bias responses in the interest of higher order ongoing task goals (Badre, 2008). However, aPFC stimulation also made participants' JOLs less sensitive to the fact that high fluency of encoding is nonpredictive of memory success. In other words, in the case of high-frequency words, participants made JOLs that were even more dependent on fluency, resulting in poorer JOL accuracy. These results are perhaps more consistent with the theory that the aPFC does not act to bias individuals toward more accurate JOLs but rather simply integrates all lower level concrete stimulus information from posterior regions (Badre, 2008), thus increasing reliance on cues that would both impair and improve JOL accuracy. However, we also found that JOL accuracy was comparably poor in all fluency conditions under sham stimulation, suggesting participants based JOLs on nondiagnostic cues in all conditions (giving higher JOLs to fluent conditions and lower JOLs to disfluent conditions). Therefore, in the current study, if the aPFC integrated all cue information regardless of diagnosticity, stimulation to this region would make JOL accuracy worse for all conditions, because participants base their JOLs on the nondiagnostic cue of fluency in all conditions. Thus, the interpretation of the role of the aPFC, namely, whether it actively biases individuals toward reliance on diagnostic cues or integrates all information from posterior regions irrespective of diagnosticity, remains unclear and further experimentation is needed.

In the context of theories proposing a hierarchical organization of pFC, the finding that the DLPFC impaired JOL accuracy for high-frequency words could also reflect increased reliance on stimulus-level properties when making JOLs: If posterior regions of pFC track lower level sensory information related to the stimulus, enhancing activity in the DLPFC may have enhanced the salience of the fluency cue, making it a more prominent basis on which participants based JOLs, resulting in poorer JOL accuracy. However, we did not find that DLPFC stimulation altered mean or median JOL ratings for any cue type, which might have been expected if stimulation enhanced the salience of cues on which JOLs are based, and further research should address whether the DLPFC plays a direct role in JOL accuracy or whether it tracks only stimulus-level properties.

Limitations

Our inferences about the roles of the aPFC and DLPFC in JOLs and encoding rest on the assumption that we are

administering focal stimulation to these precise regions using HD-tDCS. However, despite the ability of HD-tDCS to administer relatively more focal stimulation as compared with conventional tDCS (Villamar et al., 2013), it is still possible that current reached cortical regions outside the ROIs in our study (Bai, Dokos, Ho, & Loo, 2014; Bikson, Rahman, & Datta, 2012; Datta et al., 2009) and that the differences we saw between DLPFC and aPFC stimulation conditions reflect relative differences between potentially broader regions stimulated by our montages. Indeed, multimodal studies have shown that stimulation may alter activity in more distal cortical regions that are functionally connected to the sites of stimulation (Hampstead et al., 2014; Jacobson, Koslowsky, & Lavidor, 2012; Keeser et al., 2011). Nevertheless, studies have shown that HD-tDCS stimulation is relatively more focal than conventional tDCS (Kuo et al., 2013; Villamar et al., 2013), and here we show relative differences between the effects of aPFC and DLPFC stimulation, suggesting that the roles of these regions in JOL processes are at least somewhat dissociable.

Our findings are consistent with past noninvasive stimulation research showing a causal role of the aPFC in JOL accuracy (Ryals et al., 2016) and a causal role of the DLPFC in subsequent false alarms (Demeter et al., 2016). However, there are important differences between stimulation methods that distinguish our results from past research. Namely, Ryals et al. (2016) found that continuous TBS, which is thought to be inhibitory to the aPFC, improved JOL accuracy, and here, we found similar results using anodal HD-tDCS, which is thought to be excitatory. Additionally, Demeter et al. (2016) found that excitatory short TBS to the DLPFC led to decreased subsequent false alarms; using anodal HD-tDCS, which is also thought to be excitatory, we showed increased subsequent false alarms. When considering the most simplistic mechanisms of continuous TBS as being inhibitory, short TBS as being excitatory, and anodal tDCS as being excitatory, this would indicate opposing findings. The effects of brain stimulation on cognitive performance, particularly in the case of tDCS, are not well enough understood, and previous research has demonstrated that anodal stimulation can produce both improvements and deficits in performance (Gaynor & Chua, 2017; Berryhill & Jones, 2012; Jacobson et al., 2012), so it may not be surprising that the directionality of results differs from TMS-based studies. TMS protocols typically administer participant-specific amounts of power to induce neuronal firing (i.e., a percentage of motor threshold), whereas tDCS applies the same amount of low dose current to all individuals. Thus, TMS and tDCS differ in that TMS directly depolarizes or hyperpolarizes underlying neurons to the point of inducing or inhibiting an action potential, whereas tDCS is thought to induce subthreshold changes in the membrane potential (Bikson et al., 2004, 2012). One potential implication of this is that TMS alters the firing pattern for all of the neurons reached by the stimulation, whereas tDCS would only lead to altering the firing pattern for neurons involved in the task and close to threshold. A related critical difference between tDCS and TBS methods relates to the spatial resolution of tDCS: despite using a high-definition 4×1 montage in the current experiment, which has improved focality as compared with conventional 1×1 tDCS (Kuo et al., 2013), the area of stimulation remains quite large relative to the spatial focality of TMS, which is thought to be within 10–20 mm (Sparing & Mottaghy, 2008; Bohning, He, George, & Epstein, 2001). Thus, it is possible that tDCS stimulation influenced activity in a larger cortical region as compared with TMS studies of aPFC function and that activity in these surrounding prefrontal subregions contributed to the effects we found on JOL accuracy. Furthermore, because tDCS alters activity by inducing changes in subthreshold neurons, the effects of tDCS on behavior are also dependent on a variety of factors related to endogenous brain activity (Berryhill et al., 2014; Brunoni et al., 2012; Datta et al., 2012), which may vary based on anatomical differences, age, mood state, and levels of arousal. Moreover, these differences may be compounded by the fact that stimulating one area has downstream consequences for networks of regions that are functionally interconnected with the aPFC (Hampstead et al., 2014; Keeser et al., 2011). Therefore, with regard to reconciling our findings with previous stimulation studies, specifically the findings of Ryals et al. (2016) and Demeter et al. (2016), it is plausible that a combination of differences between stimulation methods, including spatial focality, neuronal populations affected, and the impact of interindividual variability, contributed to distinct network-related activity in regions functionally interconnected with the aPFC that differs between tDCS and TMS methods (Fertonani & Miniussi, 2017). We attempted to minimize the degree to which factors related to underlying brain activity influenced our results by controlling for individual differences in sex, age, head circumference, stimulation experience (e.g., itching, headache), and mood states (e.g., trouble concentrating, sleepiness) in our analyses of the effects of tDCS on memory and metamemory; however, further research is needed to test how these individual differences mediate the effects of tDCS on other forms of cognition, with an aim to better predict who will benefit from tDCS.

Another caveat is that the sham condition resulted in imperfect blinding, and participants may have been able to distinguish between sham and active HD-tDCS. However, there were no differences between the DLPFC and aPFC conditions in terms of participants' guesses about whether they were receiving active or sham stimulation. Thus, our strongest results are ones that show differences between the DLPFC and aPFC conditions, namely, the increase in false alarms in the DLPFC condition and the improvements in metamemory accuracy in the aPFC condition under disfluent conditions (i.e., low-frequency and inverted words). More broadly, the imperfect blinding highlights the benefits

of using active control sites in tDCS work in addition to sham.

Conclusions

The current experiment provides evidence for a causal role of the DLPFC in subsequent false alarms and of the DLPFC and aPFC in JOL accuracy. Importantly, this is the first study to investigate how the cognitive basis of JOLs interacts with the regions of pFC thought to be involved in JOL processes. We showed that the aPFC contributes to accurate JOLs made based on word frequency and orientation by biasing individuals toward basing their subjective judgments on disfluency, which is predictive of objective memory success. Conversely, DLPFC and aPFC stimulation decreased JOL accuracy for high-frequency words, suggesting the roles of these regions in JOLs vary with the cognitive bases of the judgments. We also demonstrated that HD-tDCS may be an effective method by which to enhance metamemory accuracy, but that it may also impair accuracy, and the nature of the effect varies with the cognitive basis of subjective confidence. Specifically, stimulation made participants more aware of their encoding failures under conditions of difficult encoding, which has implications for strategic learning and behavioral interventions, such as in the domain of education, that rely on a learner's ability to accurately identify conditions that impede or facilitate successful encoding. These results provide an important contribution to our understanding of the causal roles of pFC subregions in memory and metamemory processes and how they may vary with cognitive mechanisms underlying these functions, and our findings have important clinical implications for the use of noninvasive brain stimulation to improve metamemory deficits.

Acknowledgments

We would like to thank Sandry Garcia for assisting with participant recruitment and data collection. This work was supported by a Doctoral Student Research Grant from the City University of New York, and A. M. G. was supported by the National Cancer Institute (P30 CA008748 and T32 CA009461). E. F. C. was supported by the National Institute of General Medical Sciences (SC3 GM121192).

Reprint requests should be sent to Elizabeth F. Chua, Department of Psychology, Brooklyn College of the City University of New York, 2900 Bedford Ave, Brooklyn, NY 11210, or via e-mail: echua@brooklyn.cuny.edu.

REFERENCES

Badre, D. (2008). Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. *Trends in Cognitive Sciences*, *12*, 193–200.

Bai, S., Dokos, S., Ho, K.-A., & Loo, C. (2014). A computational modelling study of transcranial direct current stimulation montages used in depression. *Neuroimage*, *87*, 332–344.

Baird, B., Cieslak, M., Smallwood, J., Grafton, S. T., & Schooler, J. W. (2015). Regional white matter variation associated with domain-specific metacognitive accuracy. *Journal of Cognitive Neuroscience*, *27*, 440–452.

Baird, B., Smallwood, J., Gorgolewski, K. J., & Margulies, D. S. (2013). Medial and lateral networks in anterior prefrontal cortex support metacognitive ability for memory and perception. *Journal of Neuroscience*, *33*, 16657–16665.

Balota, D. A., Burgess, G. C., Cortese, M. J., & Adams, D. R. (2002). The word-frequency mirror effect in young, old, and early-stage Alzheimer's disease: Evidence for two processes in episodic recognition performance. *Journal of Memory and Language*, *46*, 199–226.

Begg, I., Duft, S., Lalonde, P., Melnick, R., & Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language*, *28*, 610–632.

Benjamin, A. S. (2003). Predicting and postdicting the effects of word frequency on memory. *Memory & Cognition*, *31*, 297–305.

Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General*, *127*, 55–68.

Benjamin, A. S., & Diaz, M. (2008). Measurement of relative metamnemonic accuracy. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 73–94). New York: Psychology Press.

Berryhill, M. E., & Jones, K. T. (2012). tDCS selectively improves working memory in older adults with more education. *Neuroscience Letters*, *521*, 148–151.

Berryhill, M. E., Peterson, D. J., Jones, K. T., & Stephens, J. A. (2014). Hits and misses: Leveraging tDCS to advance cognitive research. *Frontiers in Psychology*, *5*, 800.

Besken, M., & Mulligan, N. W. (2013). Easily perceived, easily remembered? Perceptual interference produces a double dissociation between metamemory and memory performance. *Memory & Cognition*, *41*, 897–903.

Bikson, M., Inoue, M., Akiyama, H., Deans, J. K., Fox, J. E., Miyakawa, H., et al. (2004). Effects of uniform extracellular DC electric fields on excitability in rat hippocampal slices *in vitro*. *Journal of Physiology*, *557*, 175–190.

Bikson, M., Rahman, A., & Datta, A. (2012). Computational models of transcranial direct current stimulation. *Clinical EEG and Neuroscience*, *43*, 176–183.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, J. R. Pomerantz, & FABBS Foundation (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). New York: Worth Publishers.

Blumenfeld, R. S., & Ranganath, C. (2006). Dorsolateral prefrontal cortex promotes long-term memory formation through its role in working memory organization. *Journal of Neuroscience*, *26*, 916–925.

Blumenfeld, R. S., & Ranganath, C. (2007). Prefrontal cortex and long-term memory encoding: An integrative review of findings from neuropsychology and neuroimaging. *Neuroscientist*, *13*, 280–291.

Boggio, P. S., Zaghi, S., Villani, A. B., Fecteau, S., Pascual-Leone, A., & Fregni, F. (2010). Modulation of risk-taking in marijuana users by transcranial direct current stimulation (tDCS) of the dorsolateral prefrontal cortex (DLPFC). *Drug and Alcohol Dependence*, *112*, 220–225.

Bohning, D. E., He, L., George, M. S., & Epstein, C. M. (2001). Deconvolution of transcranial magnetic stimulation (TMS) maps. *Journal of Neural Transmission*, *108*, 35–52.

Braboszcz, C., & Delorme, A. (2011). Lost in thoughts: Neural markers of low alertness during mind wandering. *Neuroimage*, *54*, 3040–3047.
Brunoni, A. R., Amadera, J., Berbel, B., Volz, M. S., Rizzerio, B. G., & Fregni, F. (2011). A systematic review on reporting and assessment of adverse effects associated with transcranial direct current stimulation. *International Journal of Neuropsychopharmacology*, *14*, 1133–1145.
Brunoni, A. R., Nitsche, M. A., Bolognini, N., Bikson, M., Wagner, T., Merabet, L., et al. (2012). Clinical research with transcranial direct current stimulation (tDCS): Challenges and future directions. *Brain Stimulation*, *5*, 175–195.
Chua, E. F., & Ahmed, R. (2016). Electrical stimulation of the dorsolateral prefrontal cortex improves memory monitoring. *Neuropsychologia*, *85*, 74–79.
Chua, E. F., Ahmed, R., & Garcia, S. M. (2017). Effects of HD-tDCS on memory and metamemory for general knowledge questions that vary by difficulty. *Brain Stimulation*, *10*, 231–241.
Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*, 201–215.
Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684.
Datta, A., Bansal, V., Diaz, J., Patel, J., Reato, D., & Bikson, M. (2009). Gyri-precise head model of transcranial direct current stimulation: Improved spatial focality using a ring electrode versus conventional rectangular pad. *Brain Stimulation*, *2*, 201–207.
Datta, A., Truong, D., Minhas, P., Parra, L. C., & Bikson, M. (2012). Inter-individual variation during transcranial direct current stimulation and normalization of dose using MRI-derived computational models. *Frontiers in Psychiatry*, *3*, 91.
Davis, N. J., Gold, E., Pascual-Leone, A., & Bracewell, R. M. (2013). Challenges of proper placebo control for non-invasive brain stimulation in clinical and experimental applications. *European Journal of Neuroscience*, *38*, 2973–2977.
Demeter, E., Mirdamadi, J. L., Meehan, S. K., & Taylor, S. F. (2016). Short theta burst stimulation to left frontal cortex prior to encoding enhances subsequent recognition memory. *Cognitive, Affective, & Behavioral Neuroscience*, *16*, 724–735.
Diana, R. A., & Reder, L. M. (2006). The low-frequency encoding disadvantage: Word frequency affects processing demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 805–815.
Diemand-Yauman, C., Oppenheimer, D. M., & Vaughan, E. B. (2011). Fortune favors the bold (and the italicized): Effects of disfluency on educational outcomes. *Cognition*, *118*, 111–115.
Do Lam, A. T. A., Axmacher, N., Fell, J., Staresina, B. P., Gauggel, S., Wagner, T., et al. (2012). Monitoring the mind: The neurocognitive correlates of metamemory. *PLoS One*, *7*, e30009.
Duarte, A., Ranganath, C., & Knight, R. T. (2005). Effects of unilateral prefrontal lesions on familiarity, recollection, and source memory. *Journal of Neuroscience*, *25*, 8333–8337.
Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.
Fertonani, A., & Miniussi, C. (2017). Transcranial electrical stimulation: What we know and do not know about mechanisms. *Neuroscientist*, *23*, 109–123.
Fleming, S. M. (2017). HMeta-d: Hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. *Neuroscience of Consciousness*, *2017*, nix007.
Gandiga, P. C., Hummel, F. C., & Cohen, L. G. (2006). Transcranial DC stimulation (tDCS): A tool for double-blind sham-controlled clinical studies in brain stimulation. *Clinical Neurophysiology*, *117*, 845–850.
Gaynor, A. M., & Chua, E. F. (2017). tDCS over the prefrontal cortex alters objective but not subjective encoding. *Cognitive Neuroscience*, *8*, 156–161.
Hampstead, B. M., Brown, G. S., & Hartley, J. F. (2014). Transcranial direct current stimulation modulates activation and effective connectivity during spatial navigation. *Brain Stimulation*, *7*, 314–324.
Harrison, B. J., Pujol, J., Ortiz, H., Fornito, A., Pantelis, C., & Yücel, M. (2008). Modulation of brain resting-state networks by sad mood induction. *PLoS One*, *3*, e1794.
Henson, R. N. A., Rugg, M. D., Shallice, T., Josephs, O., & Dolan, R. J. (1999). Recollection and familiarity in recognition memory: An event-related functional magnetic resonance imaging study. *Journal of Neuroscience*, *19*, 3962–3972.
Hertzog, C., Dunlosky, J., Robinson, A. E., & Kidder, D. P. (2003). Encoding fluency is a cue used for judgments about learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 22–34.
Higham, P. A., Perfect, T. J., & Bruno, D. (2009). Investigating strength and frequency effects in recognition memory using type-2 signal detection theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 57–80.
Horvath, J. C., Carter, O., & Forte, J. D. (2014). Transcranial direct current stimulation: Five important issues we aren't discussing (but probably should be). *Frontiers in Systems Neuroscience*, *8*, 2.
Jacobson, L., Koslowsky, M., & Lavidor, M. (2012). tDCS polarity effects in motor and cognitive domains: A meta-analytical review. *Experimental Brain Research*, *216*, 1–10.
Jia, X., Li, P., Li, X., Zhang, Y., Cao, W., Cao, L., et al. (2016). The effect of word frequency on judgments of learning: Contributions of beliefs and processing fluency. *Frontiers in Psychology*, *6*, 1995.
Johnston, W. A., Dark, V. J., & Jacoby, L. L. (1985). Perceptual fluency and recognition judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 3–11.
Kao, Y.-C., Davis, E. S., & Gabrieli, J. D. E. (2005). Neural correlates of actual and predicted memory formation. *Nature Neuroscience*, *8*, 1776–1783.
Keeser, D., Meindl, T., Bor, J., Palm, U., Pogarell, O., Mulert, C., et al. (2011). Prefrontal transcranial direct current stimulation changes connectivity of resting-state networks during fMRI. *Journal of Neuroscience*, *31*, 15284–15293.
Kishiyama, M. M., & Yonelinas, A. P. (2003). Novelty effects on recollection and familiarity in recognition memory. *Memory & Cognition*, *31*, 1045–1051.
Kishiyama, M. M., Yonelinas, A. P., & Knight, R. T. (2009). Novelty enhancements in memory are dependent on lateral prefrontal cortex. *Journal of Neuroscience*, *29*, 8114–8118.
Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, *126*, 349–370.
Koriat, A. (2008). Easy comes, easy goes? The link between learning and remembering and its exploitation in metacognition. *Memory & Cognition*, *36*, 416–428.
Koriat, A., & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory and Language*, *52*, 478–492.
Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
Kuo, H.-I., Bikson, M., Datta, A., Minhas, P., Paulus, W., Kuo, M.-F., et al. (2013). Comparing cortical plasticity induced by

conventional and high-definition 4 × 1 ring tDCS: A neurophysiological study. *Brain Stimulation*, *6*, 644–648.
Learmonth, G., Thut, G., Benwell, C. S. Y., & Harvey, M. (2015). The implications of state-dependent tDCS effects in aging: Behavioural response is determined by baseline performance. *Neuropsychologia*, *74*, 108–119.
Macmillan, N. A., Rotello, C. M., & Miller, J. O. (2004). The sampling distributions of Gaussian ROC statistics. *Perception & Psychophysics*, *66*, 406–421.
Maniscalco, B., & Lau, H. (2014). Signal detection theory analysis of type 1 and type 2 data: Meta-d', response-specific meta-d', and the unequal variance SDT model. In S. M. Fleming & C. D. Frith (Eds.), *The cognitive neuroscience of metacognition* (pp. 25–66). New York: Springer-Verlag.
Marshall, L., Mölle, M., Siebner, H. R., & Born, J. (2005). Bifrontal transcranial direct current stimulation slows reaction time in a working memory task. *BMC Neuroscience*, *6*, 23.
Masson, M. E. J., & Rotello, C. M. (2009). Sources of bias in the Goodman–Kruskal gamma coefficient measure of association: Implications for studies of metacognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 509–527.
Matzen, L. E., Trumbo, M. C., Leach, R. C., & Leshikar, E. D. (2015). Effects of non-invasive brain stimulation on associative memory. *Brain Research*, *1624*, 286–296.
Mayberg, H. S., Liotti, M., Brannan, S. K., McGinnis, S., Mahurin, R. K., Jerabek, P. A., et al. (1999). Reciprocal limbic-cortical function and negative mood: Converging PET findings in depression and normal sadness. *American Journal of Psychiatry*, *156*, 675–682.
Metcalfe, J. (2002). Is study time allocated selectively to a region of proximal learning? *Journal of Experimental Psychology: General*, *131*, 349–363.
Miele, D. B., Finn, B., & Molden, D. C. (2011). Does easily learned mean easily remembered?: It depends on your beliefs about intelligence. *Psychological Science*, *22*, 320–324.
Monti, A., Cogiamanian, F., Marceglia, S., Ferrucci, R., Mameli, F., Mrakic-Sposta, S., et al. (2008). Improved naming after transcranial direct current stimulation in aphasia. *Journal of Neurology, Neurosurgery & Psychiatry*, *79*, 451–453.
Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect.". *Psychological Science*, *2*, 267–270.
Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and some new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 125–173). New York: Academic Press.
Neuling, T., Rach, S., & Herrmann, C. S. (2013). Orchestrating neuronal networks: Sustained after-effects of transcranial alternating current stimulation depend upon brain states. *Frontiers in Human Neuroscience*, *7*, 161.
O'Connell, N. E., Cossar, J., Marston, L., Wand, B. M., Bunce, D., Moseley, G. L., et al. (2012). Rethinking clinical trials of transcranial direct current stimulation: Participant and assessor blinding is inadequate at intensities of 2 mA. *PLoS One*, *7*, e47514.
Parkin, A. J., Bindschaedler, C., Harsent, L., & Metzler, C. (1996). Pathological false alarm rates following damage to the left frontal cortex. *Brain and Cognition*, *32*, 14–27.
Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*, 8–13.
Rabinowitz, J. C., Ackerman, B. P., Craik, F. I. M., & Hinchley, J. L. (1982). Aging and metamemory: The roles of relatedness and imagery. *Journal of Gerontology*, *37*, 688–695.
Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General*, *137*, 615–625.
Rudell, A. P. (1993). Frequency of word usage and perceived word difficulty: Ratings of Kučera and Francis words. *Behavior Research Methods, Instruments, & Computers*, *25*, 455–463.
Ryals, A. J., Rogers, L. M., Gross, E. Z., Polnaszek, K. L., & Voss, J. L. (2016). Associative recognition memory awareness improved by theta-burst stimulation of frontopolar cortex. *Cerebral Cortex*, *26*, 1200–1210.
Shaughnessy, J. J. (1981). Memory monitoring accuracy and modification of rehearsal strategies. *Journal of Verbal Learning and Verbal Behavior*, *20*, 216–230.
Slotnick, S. D., & Schacter, D. L. (2004). A sensory signature that distinguishes true from false memories. *Nature Neuroscience*, *7*, 664–672.
Sparing, R., & Mottaghy, F. M. (2008). Noninvasive brain stimulation with transcranial magnetic or direct current stimulation (TMS/tDCS)—From insights into human memory to therapy of its dysfunction. *Methods*, *44*, 329–337.
Sperling, R. A., Bates, J. F., Cocchiarella, A. J., Schacter, D. L., Rosen, B. R., & Albert, M. S. (2001). Encoding novel face–name associations: A functional MRI study. *Human Brain Mapping*, *14*, 129–139.
Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*, 137–149.
Sungkhasettee, V. W., Friedman, M. C., & Castel, A. D. (2011). Memory and metamemory for inverted words: Illusions of competency and desirable difficulties. *Psychonomic Bulletin & Review*, *18*, 973–978.
Toth, J. P., Daniels, K. A., & Solinger, L. A. (2011). What you know can hurt you: Effects of age and prior knowledge on the accuracy of judgments of learning. *Psychology and Aging*, *26*, 919–931.
Tremblay, S., Lepage, J.-F., Latulipe-Loiselle, A., Fregni, F., Pascual-Leone, A., & Théoret, H. (2014). The uncertain outcome of prefrontal tDCS. *Brain Stimulation*, *7*, 773–783.
Vilkki, J., Surma-aho, O., & Servo, A. (1999). Inaccurate prediction of retrieval in a face matrix learning task after right frontal lobe lesions. *Neuropsychology*, *13*, 298–305.
Villamar, M. F., Volz, M. S., Bikson, M., Datta, A., DaSilva, A. F., & Fregni, F. (2013). Technique and considerations in the use of 4 × 1 ring high-definition transcranial direct current stimulation (HD-tDCS). *Journal of Visualized Experiments*, e50309.
Wheeler, M. A., Stuss, D. T., & Tulving, E. (1995). Frontal lobe damage produces episodic memory impairment. *Journal of the International Neuropsychological Society*, *1*, 525–536.
Whittlesea, B. W. A., & Williams, L. D. (2000). The source of feelings of familiarity: The discrepancy-attribution hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 547–565.
Wilson, M. (1988). MRC psycholinguistic database: Machine-usable dictionary, version 2.00. *Behavior Research Methods, Instruments, & Computers*, *20*, 6–10.
Yang, H., Cai, Y., Liu, Q., Zhao, X., Wang, Q., Chen, C., et al. (2015). Differential neural correlates underlie judgment of learning and subsequent memory performance. *Frontiers in Psychology*, *6*, 1699.
Yonelinas, A. P., Otten, L. J., Shaw, K. N., & Rugg, M. D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *Journal of Neuroscience*, *25*, 3002–3008.
Yue, C. L., Castel, A. D., & Bjork, R. A. (2013). When disfluency is—and is not—a desirable difficulty: The influence of typeface clarity on metacognitive judgments and memory. *Memory & Cognition*, *41*, 229–241.