

**Project Title:** The Role of Different Prefrontal Areas in Visual Metacognition

**NCT Number:** NCT04263766

**Date:** 06/12/2023

## Study Protocol

Metacognition refers to the ability to self-assess the quality of one's own decisions, and it is crucial for effective decision-making. A lot of prior research has focused on revealing the neural correlates underlying human confidence judgments, and many studies have pointed to a central role for the prefrontal cortex and specifically the dorsolateral prefrontal cortex (DLPFC). However, although these studies made progress in discovering where confidence is computed in the brain, much less is known about the timing of confidence computation. A recent paper by Shekhar and Rahnev (2018) examined this question by delivering a train of three pulses of transcranial magnetic stimulation (TMS) to DLPFC at 250, 350, and 450 ms after stimulus onset in a perceptual decision-making task. From this study, they found that the TMS train of pulses decreased confidence compared to a control region (the primary somatosensory cortex) but could not determine exactly when confidence computation occurred.

To address more precisely the issue of when confidence computations occur, we here developed a protocol where we used single pulse TMS at four different times. Specifically, we delivered single pulses of TMS at 200, 300, 400, and 500 ms after stimulus onset and compared the results to TMS delivered simultaneously with stimulus onset (0 ms condition). The task for participants was to judge the orientation of a briefly presented visual stimulus and to report confidence ratings on the decisions. We delivered online TMS to DLPFC in the experimental group, and to vertex in the control group.

The main experiment comprised four sets, each consisting of five blocks with 40 trials in each block, resulting in a total of 800 trials. The possible delays for transcranial magnetic stimulation (TMS) - 0, 200, 300, 400, and 500 ms after the stimulus onset - were presented in a pseudorandom order. Within each group of five trials, each delay appeared once. The design and procedure were the same for both the DLPFC and vertex groups, except for the specific target site.

Before starting the experiment, participants underwent a behavioral training session without TMS. The training began with a high contrast value (80%) for Gabor patches and gradually decreased the contrast over subsequent blocks (the final block had a contrast value of 6%). Participants received feedback on their performance after each trial during the training. Following this, participants completed a staircasing procedure known as 3-down-1-up without feedback. This adaptive procedure estimated the appropriate contrast value for each individual participant. On average, this procedure yielded a mean contrast value of 6.64% (standard deviation = 0.96%). The contrast value obtained for each participant during this procedure was used for the rest of the experiment. At the end of the training, participants performed one practice TMS block with the same contrast level and TMS delivery as in the rest of the experiment. The practice block aimed to familiarize participants with receiving TMS while performing the task and was not included in further analyses.

During each trial of the task, participants initially fixated on a small white dot ( $0.05^\circ$  in size) at the center of the screen for 500 ms. This was followed by the presentation of a Gabor patch ( $3^\circ$  in diameter) that appeared either tilted to the right ( $45^\circ$  clockwise) or left ( $135^\circ$

counterclockwise) relative to the vertical orientation, lasting for 100 ms. The Gabor patch was superimposed on a noisy background. Participants indicated the tilt orientation of the Gabor patch while simultaneously rating their confidence on a 4-point scale (1 denoting the lowest confidence and 4 denoting the highest confidence) by pressing a single key. The four fingers of the left hand corresponded to the four confidence ratings for the left tilt response, while the four fingers of the right hand corresponded to the four confidence ratings for the right tilt response. For each hand, the index finger represented a confidence rating of 1, while the pinky finger represented a confidence rating of 4. The stimulus orientation (left/right) was randomly determined on each trial.

Two sites were selected as TMS targets: the right dorsolateral prefrontal cortex (DLPFC) and the vertex. The right DLPFC was localized based on previous studies, using the F4 electrode position in the 10-20 system used for EEG electrode placement. Consistent with other studies targeting DLPFC with TMS during perceptual decision-making tasks, the right hemisphere was chosen as it is dominant for visual processing. To determine the precise stimulation location specific to each participant, we employed the Beam F3 Location System developed by Beam and Borckardt (2009). This method allowed us to accurately identify the F4 region using skull measurements. The vertex location was determined as the midpoint between the nasion and inion.

## Statistical Analysis Plan

We conducted an analysis of confidence and metacognitive ability for each delay condition. Metacognitive ability was measured using the Mratio metric, which was developed by Maniscalco and Lau (2012). Mratio is based on signal detection theory and involves the observer's decision and confidence responses. It is calculated as the ratio of two measures: the observer's metacognitive sensitivity (meta- $d'$ ), which reflects their ability to distinguish between correct and incorrect responses, and the observer's stimulus sensitivity ( $d'$ ), which reflects their ability to differentiate between the two stimulus classes. The meta- $d'$  to  $d'$  ratio accounts for the influence of stimulus sensitivity on metacognitive performance and captures the efficiency of the observer's metacognitive processes.

To investigate the impact of TMS on confidence and metacognitive ability (Mratio), we calculated the difference between the confidence and Mratio scores for each delay condition compared to the 0-ms condition. Subsequently, we employed one-way and two-way repeated-measures ANOVAs to compare the obtained difference scores between the two TMS stimulation sites (DLPFC and vertex) and the four TMS delay conditions (200, 300, 400, and 500 ms). Independent sample t-tests were used for direct comparisons between the two TMS stimulation sites within each delay condition, while paired t-tests were used for direct comparisons between different delay conditions within a single stimulation site.