The Relationship between Cognitive and Neural Bases of Metamemory Judgments

by

Alexandra M. Gaynor

Metamemory monitoring, the process of making subjective assessments of the status of one's own memory, is crucial to guiding behavior and effective learning. Past cognitive research has shown that subjective confidence judgments are inferential in nature, and based on cues available at the time of the judgment. When confidence is based on cues that are related to objective memory performance, metamemory accuracy is high. However, past studies have shown that metamemory monitoring tends to be inaccurate because individuals base their confidence on information that is not predictive of memory success, such as the fluency with which items were encoded during study, or invalid information about task difficulty from external sources. Brain research has lagged behind cognitive research in establishing the neural bases of metamemory monitoring, but there is evidence that the prefrontal cortex contributes to Judgments of Learning (JOLs), subjective confidence judgments made during encoding about the ability to retrieve information at later test. Retrospective Confidence Judgments (RCJs), which are made at retrieval and reflect one's confidence in his or her response, have been associated with activity in the prefrontal, parietal, and temporal lobes. However, there is evidence that the roles of each of these regions in metamemory monitoring may vary with the information on which confidence is based, and here we present two experiments which tested how the neural mechanisms underlying JOLs and RCJs interact with the cognitive bases of the judgments. Experiment 1 used transcranial direct current stimulation (tDCS) to directly manipulate activity in the dorsolateral prefrontal cortex (DLPFC) and anterior prefrontal cortex (aPFC) during a JOL task in which subjects studied words that varied in their fluency at encoding. Results showed that DLPFC stimulation impaired encoding, and aPFC stimulation enhanced JOL accuracy for disfluent encoding conditions. Conversely, DLPFC and aPFC stimulation decreased JOL accuracy for fluently encoded conditions, suggesting the causal roles of the aPFC and DLPFC in JOL accuracy vary depending on whether JOLs are based on the cue of fluency or disfluency.

Experiment 2 used fMRI to explore the neural correlates of RCJs and how confidence-related activity varies when RCJs are based on internal memory cues as compared to external cues about question difficulty during a semantic recognition task. Results showed that as compared to recognition, the process of making RCJs engaged regions consistent with those identified using episodic memory tasks. As compared to cued trials, uncued trials produced activity in lateral and medial parietal, and inferior temporal regions, which may reflect self-referential processing or uncertainty. Activity in the parietal lobe was greater for invalid than valid cues during hard questions, while occipital and subcortical regions showed greater activity during valid than invalid cueing for easy questions, suggesting confidence-related activity varies with both cue validity and actual question difficulty. Lastly, activity in bilateral medial temporal lobes covaried with the degree to which subjects incorporated invalid cues into their confidence judgments for correctly answered hard questions, which we speculate may reflect integration of multiple sources of information during RCJs. Taken together, these experiments suggest that the roles of brain regions involved in metamemory processes differ based on the information on which individuals base their confidence assessments.