A neurocognitive model for third-party norm enforcement

Joshua Buckholtz
Harvard University

Third-party punishment of social norm violations is thought to be crucial for the development and stability of human cooperation. Brain imaging and brain stimulation suggests that dorsolateral prefrontal cortex function is critical for a range of norm-based social decisions, including third-party punishment. However, the nature of DLPFC information processing in norm-based behavior remains elusive. I will present evidence from several recent fMRI and brain stimulation studies that DLPFC integrates culpability signals (arising in TPJ) with harm signals (arising in amygdala), and uses this integrated representation to select actions during third-party punishment. These data suggest a specific, causal network mechanism underlying human norm enforcement behavior. Intriguingly, the engagement of this network is evident across norm-enforcement contexts (explicit legal norms and implicit social cooperative norms) and across groups that vary according in norm enforcement experience (i.e. trial court judges and lay volunteers). Together, these studies suggest that DLPFC supports norm enforcement through its ability to receive and integrate representations of decision-relevant variables that are generated at other cortical and subcortical sites.