Neuroscience and Biobehavioral Reviews 71 (2016) 101–111

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Review article

Neural signatures of social conformity: A coordinate-based activation likelihood estimation meta-analysis of functional brain imaging studies

Haiyan Wu[a], Yi Luo[b], Chunliang Feng[b,*]

[a] Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, 100101, Beijing, China
[b] State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, 100875, Beijing, China

A R T I C L E I N F O

Article history:
Received 28 June 2016
Received in revised form 28 August 2016
Accepted 30 August 2016
Available online 31 August 2016

Keywords:
Social conformity
Norm violations
Activation likelihood estimation
Meta-analysis
Ultimatum game
VS (Ventral striatum)
Dorsal pMFC (dorsal posterior medial frontal cortex)
AI (Anterior insula)

A B S T R A C T

People often align their behaviors with group opinions, known as social conformity. Many neuroscience studies have explored the neuropsychological mechanisms underlying social conformity. Here we employed a coordinate-based meta-analysis on neuroimaging studies of social conformity with the purpose to reveal the convergence of the underlying neural architecture. We identified a convergence of reported activation foci in regions associated with normative decision-making, including ventral striatum (VS), dorsal posterior medial frontal cortex (dorsal pMFC), and anterior insula (AI). Specifically, consistent deactivation of VS and activation of dorsal pMFC and AI are identified when people's responses deviate from group opinions. In addition, the deviation-related responses in dorsal pMFC predict people's conforming behavioral adjustments. These are consistent with current models that disagreement with others might evoke "error" signals, cognitive imbalance, and/or aversive feelings, which are plausibly detected in these brain regions as control signals to facilitate subsequent conforming behaviors. Finally, group opinions result in altered neural correlates of valuation, manifested as stronger responses of VS to stimuli endorsed than disliked by others.

Contents

* Corresponding author.
E-mail address: chunliang.feng@gmail.com (C. Feng).

http://dx.doi.org/10.1016/j.neubiorev.2016.08.038

1. Introduction

Human preferences, judgments, and attitudes are highly sensitive to social influence, such that people often adjust their behaviors to match the responses of others (Cialdini and Goldstein, 2004; Wood, 2000). One of the most basic and best-known forms of social influence is the human conformity behavior initially reported in the seminal work of Solomon Asch (Asch, 1951, 1956), wherein participants were asked to judge relative lengths of lines alone or in the presence of confederates posing as other subjects. Participants invariably gave the correct answer when they performed the task alone. Importantly, however, about 37% of participants conformed to the erroneous judgments of confederates in this simple task (Asch, 1951, 1956). Since these initial findings, social conformity has been identified in various tasks (e.g., Alquist et al., 2013; Gabbert et al., 2003; Koban and Wager, 2016). For instance, when there is a discrepancy between one's own and group's preferences or memories, people often subsequently change their responses to the same items to decrease the discrepancy (Campbell-Meiklejohn et al., 2010; Edelson et al., 2011; Izuma and Adolphs, 2013; Klucharev et al., 2009b). Those behavioral adjustments induced by normative opinions have been thought to be driven by two separate and interacting motivations (Allen, 1965; Kelman, 1961; Peterson et al., 1985): (i) mere public compliance, i.e., agreeing with others on the surface while maintaining one's intrinsic attitudes (e.g., Berns et al., 2010); or (ii) private acceptance, i.e., internalizing the preferences, judgments, and attitudes of others (e.g., Huang et al., 2014; Nook and Zaki, 2015).

Building on the extensive behavioral research in the social psychology literature, the past decade has witnessed a surge of interest in unveiling the neural mechanisms of social influence (e.g., Berns et al., 2005; Campbell-Meiklejohn et al., 2012a; Chen et al., 2012; Chua et al., 2011; Edelson et al., 2011; Izuma and Adolphs, 2013; Klucharev et al., 2009b; Klucharev et al., 2008; Shestakova et al., 2013). The neuroscientific approach provides additional level of evidence in validating psychological theories of social influence. For instance, as the first study to examine the neural mechanisms underlying social conformity, Berns et al. (2005) demonstrated that human conforming behaviors were paralleled by altered perceptual representations of visual stimuli in an occipital–parietal network. Likewise, social influence leads to long-lasting alterations in people's memory via modifying neural mnemonic representations in the hippocampus and amygdala (Edelson et al., 2011; Edelson et al., 2014). Last but not least, many studies have indicated that the opinions of others change people's preferences and alter neural representations of value assigned to stimuli, manifested as the modulated engagement of ventral striatum (VS) and orbital frontal cortex (OFC) (Campbell-Meiklejohn et al., 2010; Charpentier et al., 2014; Zaki et al., 2011). These findings together suggest a possibility that people internalize judgments and preferences of other people; and therefore, are thought to support the account of private acceptance (but see Berns et al., 2010).

Furthermore, a plethora of neuroscientific studies on social conformity has examined the neural responses to the discrepancy between one's own and group opinions. In particular, the consensus between oneself and others in judgments recruits brain activations of VS that plays an important role in reward-driven behaviors (Campbell-Meiklejohn et al., 2010; Klucharev et al., 2009b). In contrast, the disagreement between oneself and others induces neural activations in the dorsal posterior medial frontal cortex (dorsal pMFC, comprising dorsal anterior cingulate cortex, posterior medial frontal cortex, and supplementary motor area) and anterior insula (AI) (Izuma and Adolphs, 2013; Klucharev et al., 2009b) that are implicated in encoding negative emotions (Corradi-Dell'Acqua et al., 2016; Lamm et al., 2011; Phan et al., 2002) and monitoring conflicts/errors (Garrison et al., 2013; Ridderinkhof et al., 2004). Notably, disagreement-dependent neural responses of these regions are predictive of people's subsequent decisions to conform group opinions (Berns et al., 2010; Burke et al., 2010; Campbell-Meiklejohn et al., 2010; Huber et al., 2015; Klucharev et al., 2009b; Nook and Zaki, 2015; Prehn et al., 2014).

The involvement of dorsal pMFC and AI in detecting first-person experience of deviating from group norms in social conformity tasks complements previous observations that these regions are engaged by norm violations conducted by another person (Buckholtz and Marois, 2012; Rilling et al., 2008; Sanfey et al., 2003; Strobel et al., 2011). For instance, in the Ultimatum Game (UG), the first player (the Proposer) proposes how to divide the money; and the second player (the Responder) decides to accept (both get paid accordingly) or reject (neither gets paid) this proposal (Güth et al., 1982). In UG, the Responder usually compares the Proposer's decisions with a fairness norm (e.g., equality) and "corrects" deviations from social norms by rejecting the unfair proposals (Xiang et al., 2013). As such, the engagement of dorsal pMFC and AI has also been consistently identified when another person's behaviors deviate from social norms (Gabay et al., 2014; Sanfey et al., 2003). The common engagement of these regions in detecting norm violations independent of agents suggests a "generic" neural system consisting of dorsal pMFC and AI for detecting deviations from group norms to facilitate behavioral adjustments in line with normative opinions (see also Montague and Lohrenz, 2007; Tomlin et al., 2013; Xiang et al., 2013).

There are many explanations regarding specific functions of brain regions involved in the social conformity. For instance, the reinforcement learning (RL) account holds that the involvement of VS, dorsal pMFC, and AI reflects the detecting of general prediction errors (i.e., the differences between outcomes and expectations) that play a crucial role in guiding people's adaptive behaviors (Campbell-Meiklejohn et al., 2010; Klucharev et al., 2009b; Klucharev et al., 2011; Shestakova et al., 2013). In this regard, the consensus between oneself and others might be experienced as a rewarding outcome (Campbell-Meiklejohn et al., 2010; Nook and Zaki, 2015). In contrast, the conflicts between one's own responses and group norms might be detected as a negative prediction error which calls for the need to correct deviance from norms, i.e., aligning one's responses with normative opinions (see also Klucharev et al., 2009b; Montague and Lohrenz, 2007). Further, a "cognitive balance" account posits that the discrepancy between one's own and group opinions might be represented separately from general prediction errors (Izuma, 2013; Izuma and Adolphs, 2013). According to this account, deviations from normative opinions might specifically engage a subset of neurons in the dorsal pMFC, rather than share identical neuronal populations with general error signals in RL (Izuma and Adolphs, 2013). Lastly, some researchers have interpreted the engagement of dorsal pMFC and AI as physiological arousal and negative affective states in response to disagreement with group opinions (Berns et al., 2010). These interpretations are not necessary to be mutually exclusive, and while a meta-analysis will not allow for directly testing these accounts; we will discuss our findings in light of these models.

In this study, we employed a coordinate-based meta-analysis on fMRI studies utilizing conformity-related paradigms with the goal to identify regions most robustly involved in the following aspects of social conformity, which are often interested in the current literature: (i) the influence of group opinion on neural representation of subjective values assigned to stimuli in value-based tasks; (ii) the neural basis of agreement and disagreement between one's own and group opinions; and (iii) the associations between disagreement-dependent neural responses and conformity behaviors. Finally, by assessing correspondence across disagreement-related contrasts in social conformity tasks and unfairness-related contrasts in UG, the current meta-analysis