Editor's Note: These short, critical reviews of recent papers in the Journal, written exclusively by graduate students or postdoctoral fellows, are intended to summarize the important findings of the paper and provide additional insight and commentary. For more information on the format and purpose of the Journal Club, please see http://www.jneurosci.org/misc/ifa_features.shtml

Dissociating Guilt- and Inequity-Aversion in Cooperation and Norm Compliance

Hongbo Yu[1], Bo Shen[1], Yunlu Yin[1], Philip R. Blue[1], and Luke J. Chang[2]

[1]Center for Brain and Cognitive Sciences and Department of Psychology, Peking University, Beijing 100871, China, and [2]Department of Psychology and Neuroscience, University of Colorado, Boulder, Colorado 80309

Review of Nihonsugi et al.

Social norms provide a set of expectations regarding context-specific appropriate behavior that aids in navigating social environments (Bicchieri, 2006). Classic studies have demonstrated that norms vary across cultures (Henrich et al., 2001) and there are likely differing motivations for individuals to comply with these norms. For example, one motivation, consequentialism, emphasizes the outcome of an action as the sole measure of its moral worth (Mill, 1861/1998). From this philosophical perspective, one may avoid violating social norms simply because unfair and inequitable outcomes are bad for the greater good (e.g., distributional preferences). Alternatively, according to sentimentalism (Smith, 1759/2002), empathy with others "constitutes the moral approval. . . for agents and/or their actions" (Slote, 2010). This framework argues that people are motivated to comply with norms to avoid suffering from harming another as a result of violating the norms (e.g., guilt-aversion).

In reality, these two motivations are likely complementary and each may independently contribute to social decisions with their relative weights varying across individuals and contexts. Unfortunately, the majority of the research that uses social bargaining games to study social decision-making has been unable to effectively dissociate these two distinct motivations. This is likely a consequence of a peculiar convention in bargaining experiments to neither measure nor manipulate individuals' expectations. Thus, it has been unclear how much participants are motivated by distributional preferences (i.e., inequity-aversion) compared with disappointing a relationship partner (i.e., guilt-aversion). Fortunately, there has recently been a growing trend to both measure (Chang et al., 2011; Chang and Sanfey, 2013) and manipulate (Xiang et al., 2013) agents' expectations.

In a recent study published in The Journal of Neuroscience, Nihonsugi et al. (2015) provided an important theoretical advance to dissociate the inequity- and guilt-aversion motivations in human norm compliance and identify the brain bases for each motivation. The experimenters used a modified trust game (Charness and Dufwenberg, 2006) in which participants initially decided as an investor whether or not to invest their endowment with an anonymous trustee and reported their belief about the likelihood of the trustee reciprocating. Participants then played the role of the trustee with multiple anonymous investors while undergoing fMRI. For each trial, trustees were given information about the investor's expectation and also the payoffs each player would receive based on their decision to cooperate or defect. For example, if the trustee chose Cooperate, then the investor might receive ¥780 and the trustee ¥650; if the trustee chose Defect, then the investor could receive ¥220 and the trustee ¥910. Though the actual investors' expectations and decisions were predetermined by the experimenters, the trustees were led to believe that they were playing with real agents and were paid proportional to their payoffs in the game at the end of the experiment.

Participants' motivations in the game were inferred based on how much they considered their partners' expectations (e.g., guilt-aversion) and discrepancies between each player's payoffs (e.g., inequity-aversion) when making their decision to cooperate or defect. The basic framework for how these motivations were modeled was based on expected utility theory, which assumes that participants make decisions that maximize their expected payoff. Here, payoffs could be material (based on the amount of money the trustee receives) or psychological (based on concern for the investor's welfare) (Fehr and Camerer, 2007). The authors specifically compared psychological payoffs arising from inequitable distributional outcomes (i.e., the absolute differ-

Received March 29, 2015; revised April 30, 2015; accepted May 5, 2015.

H.Y, B.S, Y.Y, and P.R.B. are supported by grants awarded to Prof. Xiaolin Zhou from the Natural Science Foundation of China Grants 91232708, 31170972) and the National Basic Research Program of China.

ence between the two players' payoffs) (Fehr and Schmidt, 1999) and feelings of guilt, which arose from disappointing a relationship partner by making a decision that resulted in the investor receiving a smaller payoff than he/she expected (i.e., the amount of money that the investor would have received had the trustee chosen to cooperate multiplied by the investor's estimated probability of the trustee's cooperation) (

consequentialism and sentimentalism considerations independently affect norm compliance and cooperation. Moreover, these motivations appear to be encoded in separate brain circuits. We believe that combining formal mathematical modeling, neuroscientific techniques, and social psychological theories will continue to further our insight into the material basis of our social nature.

References

Battigalli P, Dufwenberg M (2007) Guilt in games. Am Econ Rev 97:170–176. CrossRef

Bicchieri C (2006) The grammar of society: the nature and dynamics of social norms. Cambridge: Cambridge UP.

Chang LJ, Sanfey AG (2013) Great expectations: neural computations underlying the use of social norms in decision-making. Soc Cogn Affect Neurosci 8:277–284. CrossRef Medline

Chang LJ, Smith A, Dufwenberg M, Sanfey AG (2011) Triangulating the neural, psychological, and economic bases of guilt aversion. Neuron 70:560–572. CrossRef Medline

Charness G, Dufwenberg M (2006) Promises and partnership. Econometrica 74:1579–1601. CrossRef

Fehr E, Camerer CF (2007) Social neuroeconomics: the neural circuitry of social preferences. Trends Cogn Sci 11:419–427. CrossRef Medline

Fehr E, Fischbacher U (2004) Third-party punishment and social norms. Evol Hum Behav 25:63–87. CrossRef

Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. Q J Econ 114:817–868. CrossRef

Henrich J, Boyd R, Bowles S, Camerer C, Fehr E, Gintis H, McElreath R (2001) In search of Homo economicus: behavioral experiments in 15 small-scale societies. Am Econ Rev 91:73–78. CrossRef

Knoch D, Schneider F, Schunk D, Hohmann M, Fehr E (2009) Disrupting the prefrontal cortex diminishes the human ability to build a good reputation. Proc Natl Acad Sci U S A 106:20895–20899. CrossRef Medline

Koban L, Corradi-Dell'Acqua C, Vuilleumier P (2013) Integration of error agency and representation of others' pain in the anterior insula. J Cogn Neurosci 25:258–272. CrossRef Medline