Guilty by association: How group-based (collective) guilt arises in the brain

Zhiai Li [a,b,1], Hongbo Yu [b,c,1], Yongdi Zhou [a], Tobias Kalenscher [d], Xiaolin Zhou [b,e,f,g,*]

[a] b

1. Introduction

Guilt is viewed as "the emotion most essential to the development of conscience and moral behavior" in field of psychology (Izard, 1991) and as the "internalized voice of moral authority" in field of philosophy (Griswold, 2007). People feel guilty when they realize that they are responsible for an action or omission that violates moral norms or mutual expectations that they accept as binding (i.e., personal guilt) (Baumeister et al., 1994, 1995; Tangney and Dearing, 2003; Taylor, 1985). Guilt can also be encountered in inter-group interactions (Halperin and Schori-Eyal, 2019; Vollberg and Cikara, 2018): individuals may feel guilty for transgressions committed by members of social groups they identify as in-group, even when he/she is not directly responsible for these transgressions. However, the psychological and neural basis of group-based (collective) guilt and its relation to personal guilt are poorly understood.

As individuals rarely engage in social interactions without social identity or association (Mesquita et al., 2016; Tajfel and Turner, 1986), social emotions arising from such interactions are often tainted by group identity and inter-group appraisals (Mackie et al., 2008; Smith and Mackie, 2015). Well-known cases of group-based or collective social

1998; McGarty et al., 2005). These studies demonstrated that group-based guilt results from the acceptance of in-group responsibility

northeastern region in mid-January, and they were required to rank-order 10 items salvaged from the broken plane (a lighter, a chocolate bar, a gun, newspaper, etc.) according to their importance for survival. The 3 individuals needed to discuss the problem together and reach a single consensus ranking within 6 min. As a manipulation check, we asked the participants to complete a scale of psychological distance to their group (a modified version of Inclusion of Others; Aron et al., 1992) and a 6-item questionnaire of group identity (example items: “How much do you identify with the Yellow group?” and “To what extent do you feel strong ties with the Yellow group?“;

upon arrival at the laboratory; 2) at the end of each estimation-failure

Our rationale for training the classifier on personal guilt is that we take personal guilt as a prototypical species of guilt that exemplifies the core cognitive-affective processes present in a family of emotions that fall into the category of guilt (e.g., survivor guilt, group-based guilt, guilt for failure in personal goals and so on) (Deigh, 1999; Schoeman, 1987; Shaver et al., 1987). The reason that those tokens of emotion are labeled as "guilt", both by emotion researchers and in everyday discourse, is because they share those core cognitive-affective processes exemplified by personal guilt. In that sense, we treat group-based guilt as a variant on the theme of personal guilt, therefore logically it makes more sense to train the classifi

Table 2
Behavioral results in Experiments 1 and 2.

Item	In-group_Commit	Out-group_Commit	In-group_Observe	Out-group_Observe	Interaction T/F

3.3. Brain activations associated with personal and group-based guilt

Here we presented the brain activation patterns revealed by the contrasts hypothesized to reflect group-based and personal guilt, respectively. The activation patterns corresponding to the main effects of Agency and Group can be found in Supplementary Neuroimaging Results.

To identify brain regions associated with group-based guilt, we focused on brain responses associated with the outcome feedback of dot estimation. We defined the critical contrast "In-group_ Observe > Out-group_ Observe", which corresponds to the effect of group-based guilt. This contrast revealed activations in anterior middle cingulate cortex (aMCC; MNI coordinates = [6, 26, 28]; k = 85 voxels) and right anterior insula (AI; MNI coordinates = [27, 20, −11]; k = 78 voxels) (Fig. 4A). aMCC and right AI have been consistently implicated in imagining and experiencing personal guilt (Chang et al., 2011; Yu et al., 2014) and negative self-evaluation in social contexts (Immordino-Yang et al., 2009; Kédia et al., 2008; Koban et al., 2013; Sanfey et al., 2003; Zaki et al., 2007). To illustrate the activation patterns, we extracted the regional parameter estimates from 27 voxels around the peak coordinates at aMCC and right AI. The parameter estimates extracted from aMCC (Fig. 4B) and right AI (Fig. 4C) exhibited a pattern similar to the pattern of monetary allocation. Moreover, aMCC activation difference (In-group_ Observe > Out-group_ Observe) positively correlated with the post-scan guilt rating difference between these two conditions ($r = 0.45$, $p = 0.011$), indicating that the aMCC was involved in the processing of group-based guilt.

We next examined whether group-based guilt shared a similar neurocognitive process with personal guilt. In the current study, 'personal guilt' was defined by the contrast 'Out-group_ Commit > Out-group_ Observe'. This contrast, while keeping the impact of group membership to its minimal, captured the difference in the participants' causal contribution to the transgression, thereby reflecting neural processing of personal guilt. Replicating previous neuroimaging findings about personal guilt (Koban et al., 2013; Yu et al., 2014), this contrast (Out-group_ Commit > Out-group_ Observe) revealed the activations in aMCC (MNI coordinates

badly would feel jointly responsible for the bad behaviors of the group. According to social identity theory (Tajfel and Turner, 1986), when

difficult in the current study to dissociate the neural processes underlying empathy and guilt. Based on the data we have, we cannot rule out the possibility that the observed effects can be explained by empathy. In fact, the observed effects in self-reported guilt and allocation could be mediated by empathy. Future studies that include direct or implicit measures of empathy are needed to empirically test this hypothesis. That being said, compared with previous studies using scenario-based task, our design has already improved in terms of controlling confounding factors such as victims' harm across critical conditions.

It might be argued that if group-based guilt is elicited by observing other group members' transgressive behaviors, then it should rely on the "Theory of Mind (ToM)" processes. We did not observe stronger activations of the so-called ToM network (e.g., temporoparietal junction, dorsomedial prefrontal cortex, etc.) in the group-based guilt condition compared with the personal guilt condition. In theory, one can experience vicarious emotion in at least two ways: one can either 'inherit' the cognitive antecedents from others and generate their own emotions based on the shared cognitive antecedents (Lickel et al., 2005), or they can directly feel the emotion that the others express, without sharing or even knowing the cognitive antecedents leading to the others' emotion (e.g., it is not uncommon that we laugh when watching a group of strangers laughing, even when we do not know why the group is laughing in the first place). The second way, also known as emotional contagion (Hatfield et al., 1993), has been shown to activate ToM related brain areas (Nummenmaa et al., 2008; Melchers et al., 2015; Müller-Pinzler et al., 2016). In this study, the way we induced group-based guilt has made it conceptually closer to the first type of vicarious emotion: 1) the participants could not see their in-group members' face when they observed the latter causing pain to the victim, neither did they know whether their in-group members feel guilty at all, therefore it was unlikely that they discerned guilt from the in-group members and took on that feelings themselves; 2) our behavioral data did show that the participants 'inherited' their in-group members' responsibility (a core cognitive antecedent of guilt) in the transgression, which was predictive of their group-based guilt and reparation (Fig. 3D). Therefore, it is not surprising that the guilt-related rather than the ToM-related network played a more important role in this study. It is an empirical question as to whether and how watching a guilt display by an in-group member can contribute to group-based guilt over and above the vicarious processing of the in-group member's responsibility.

Note that it is beyond the scope of this study to empirically discriminate guilt from other related emotions, such as regret and shame. While it is still under debate in emotion science whether and how emotions should be categorized (Griffiths, 2004; Lindquist and Barrett, 2012; Satpute et al., 2016;

moral emotion and lead to different behavioral patterns. The increasing

Bastin, C., Harrison, B.J., Davey, C.G., Moll, J., Whittle, S., 2016. Feelings of shame, embarrassment and guilt and their neural correlates: a systematic review. Neurosci. Biobehav. Rev. 71, 455–471. https://doi.org/10.1016/j.neubiorev.2016.09.019.

Baayen, R.H., Davidson, D.J., Bates, D.M., 2008. Mixed-effects modeling with crossed random effects for subjects and items. J. Mem. Lang. 59, 390–412. https://doi.org/10.1016/j.jml.2007.12.005.

Baron, M., 1988. Remorse and agent-regret. Midwest Stud. Philos. 13, 259–281. https://doi.org/10.1111/j.1475-4975.1988.tb00126.x.

Barr, D.J., Levy, R., Scheepers, C., Tily, H.J., 2013. Random effects structure for confirmatory hypothesis testing: keep it maximal. J. Mem. Lang. 68, 255–278. https://doi.org/10.1016/j.jml.2012.11.001.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2014. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67, 1–48. https://doi.org/10.18637/jss.v067.i01.

Baumeister, R.F., Stillwell, A.M., Heatherton, T.F., 1994. Guilt: an interpersonal approach. Psychol. Bull. 115, 243–267. https://doi.org/10.1037/0033-2909.115.2.243.

Baumeister, R.F., Stillwell, A.M., Heatherton, T.F., 1995. Personal narratives about guilt: role in action control and interpersonal relationships. Basic Appl. Soc. Psychol. 17, 173–198. https://doi.org/10.1207/s15324834basp1701.

Branscombe, N.R., Slugoski, B., Kappen, D.M., 2004. Collective Guilt: what it Is and what it Is Not. Cambridge University Press., Cambridge.

Brown, R., González, R., Zagefka, H., Manzi, J., Čehajić, S., 2008. Nuestra culpa: collective guilt and shame as predictors of reparation for historical wrongdoing. J. Personal. Soc. Psychol. 94, 75–90. https://doi.org/10.1037/0022-3514.94.1.75.