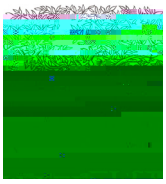
NeuroImage 244 (2021) 118631



Contents lists available at ScienceDirect

NeuroImage

. a a : . .c / ca / . a

2003; Gassin and Elizabeth, 1998). Several studies suggested that forgiveness reduces guilt (McNulty, 2010, 2011) and blame increases it (Kubany and Watson, 2003; Parkinson and Illingworth, 2009), whereas others observed the opposite e ects, namely forgiveness enhances guilt (Wallace et al., 2008), and blame reduces guilt and even induces anger [illegible]

Fig. 1. Procedures of the interactive game. In each round, after being paired with a same sex anonymous co-player, the participant would see a picture of dots for 1.5 s and estimate the number of dots quickly by choosing one of the four numbers presented on the screen within 3 s (i.e., Dot estimation). After that the correctness of the estimation was revealed. If the estimation was correct, the current trial was terminated and the game entered the next round; otherwise, the co-player (the victim) in the current round would re-

harm and Low harm conditions respectively. The other 8 trials with attitude ratings of "-2," "-1," "1," and "2"

X.

the more conventional searchlight approach, such as less computationally demanding and higher homogeneity with functional neuroanatomy (Chang et al., 2021; Craddock et al., 2012; van Baar et al., 2019), and has been proven e cient in multivariate based analysis (Chang et al., 2021; van Baar et al., 2019). Next, to identify parcels contributing to reactive guilt processing, for each parcel, we applied SVM (Friedman et al., 2001; Wager et al., 2013) to train a multivariate pattern classifier discriminating High guilt *vs.* Low guilt groups in Attitude phase (Chang et al., 2015; Wager et al., 2013; Woo et al., 2014). The same set of analyses was conducted

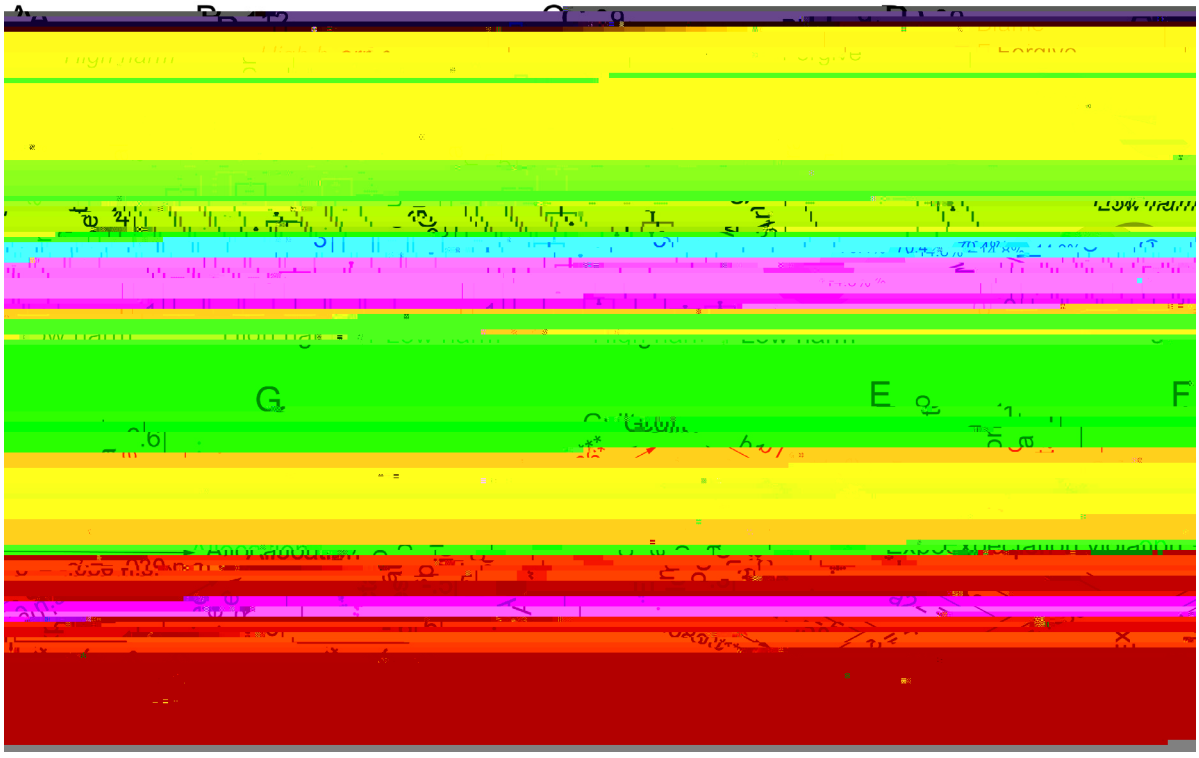


Fig. 2. Behavioral results. (A - D) Participants' pre-experiment prediction of the co-player's attitudes in High and Low harm conditions (A), amount of

Table 1
Descriptive statistics for behavioral results.

Experiment	Variable	High harm			Low harm		
		Blame	Forgive	Neutral	Blame	Forgive	Neutral
fMRI	Monetary allocation	5.08 ± 0.42	5.77 ± 0.29	5.30 ± 0.30	3.78 ± 0.34	5.50 ± 0.32	4.84 ± 0.31
	Guilt rating	4.41 ± 0.29	4.93 ± 0.37	3.74 ± 0.28	2.15 ± 0.25	3.48 ± 0.30	2.93 ± 0.28
	Anger rating	2.33 ± 0.27	1.04 ± 0.04	1.44 ± 0.16	4.11 ± 0.28	1.15 ± 0.09	1.22 ± 0.10
Behavioral Replication	Monetary allocation	2.43 ± 0.29	3.71 ± 0.38	2.86 ± 0.32	1.64 ± 0.27	3.24 ± 0.37	2.56 ± 0.29
	Guilt rating	4.47 ± 0.27	4.87 ± 0.29	3.17 ±			

Fig. 3. Whole-brain multivariate pattern analyses for reactive guilt and anger in Attitude phase. (A) Whole-brain multivariate patterns discriminating High *vs.* Low guilt conditions and High *vs.* Low anger conditions in Attitude phase. (B) Receiver operating characteristic curves (ROCs) for within-emotion and cross-emotion classifications. Orange solid, cross-validations for High *vs.* Low guilt conditions in Attitude phase; orange dash, using High *vs.* Low guilt pattern to predict High *vs.* Low anger conditions; red solid, cross-validations for High *vs.* Low anger conditions in Attitude phase; red dash, using High *vs.* Low anger pattern to predict High *vs.* Low guilt conditions. (C and D) Pattern expression values of reactive guilt and reactive anger in the six conditions

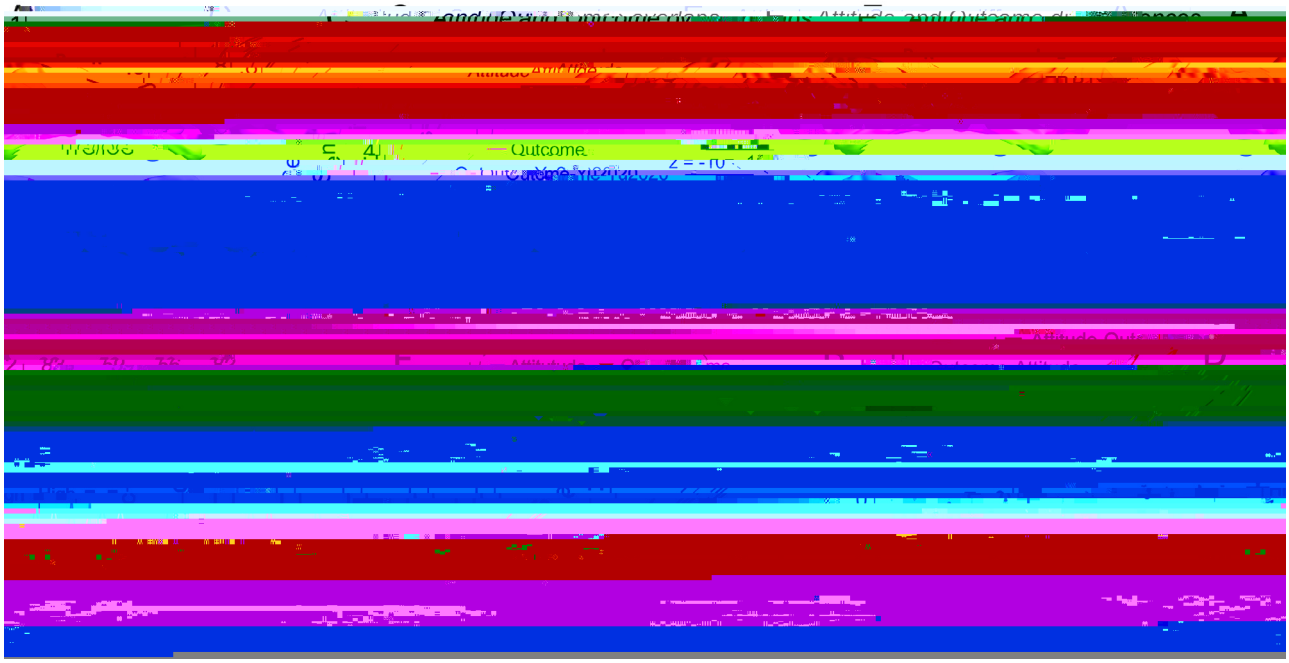


Fig. 6. Shared and di erential neural representations for reactive guilt in Attitude phase and non-reactive guilt in Outcome phase. (A) ROCs for the two-choice forced-alternative accuracies for within-study and cross-study classifications using the data of Outcome phase in the current study and of Yu et al.

tions of sample size and statistical power confirmed that our conclusions would not change if the sample size increased (see *Supplementary Materials* and Fig. S3, I-K). These results suggested that after the co-player's attitude feedback, the neural representation of guilt (i.e., reactive guilt) might di er from the representation of guilt at whole-brain level when the co-player's attitude is not involved (i.e., non-reactive guilt).

As a supplementary analysis, we examined whether the representation of reactive anger in Attitude phase could be distinguished by the multivariate pattern of anger identified in a previous meta-analysis (Wager et al.,