Identity Fusion, Outgroups, and Sacrifice: A Cross-Cultural Test
Supplementary Materials

Benjamin Grant Purzycki[a]* and Martin Lang[b,c]

[a]Max Planck Institute for Evolutionary Anthropology, Germany
[b]Harvard University, United States
[c]Masaryk University, Czech Republic

*contact author email: benjamin_purzycki@eva.mpg.de

Contents

1. Background

Over the past decade, researchers have devoted increased attention to the theory of identity fusion, which proposes that the psychological state of identity fusion with a group mediates individuals' costly sacrifices on behalf of those groups (Swann, Jetten, Gómez, Whitehouse, & Bastian, 2012). Based on the theory of self-verification (Swann, 1983; Swann & Hill, 1982), identity fusion with a group has been shown to predict an array of self-reported pro-group behaviors. For example, Swann et al. (Swann, Gómez, Seyle, Morales, & Huici, 2009) have shown that participants who select completely overlapping circles to represent their and their country's identities (fused individuals) expressed higher willingness to fight and die for their country. Interestingly, while the activation of personal identities had no effect on non-fused individuals, it increased the willingness to fight and die for the fused individuals to the same extent as activating social identities. The authors interpreted these results as a support for a functional equivalency of personal and social identities in fused individuals (Swann et al., 2009). In a subsequent study, Swann et al. (Swann, Gómez, Dovidio, Hart, & Jetten, 2010) showed that the pictorial indicator of fusion predicted individual Spaniards' willingness to sacrifice their lives to save five other Spaniards. This effect extended even to five other Europeans, but not to Americans. Moreover, increasing agency through various manipulation of arousal further increased pro-group behavior of fused individuals (Swann, Gómez, Huici, Morales, & Hixon, 2010).

Improving upon the pictorial measure that artificially divided participants into a binary category of fused/non-fused, Gómez at al. (2011) created a seven-item verbal measure of identity fusion showing that this continuous measure predicts participants' self-reported willingness to fight and die for their country and self-reported willingness to sacrifice one's life to save five other countrymen. Importantly, the relationship between the verbal measure of identity fusion and willingness to self-sacrifice was not mediated by decreased desire for self-preservation in fused participants, but rather by increased devotion to the group (Swann, Gómez, et al., 2014). The same authors have also shown that under time pressure, the psychological state of identity fusion exerts even stronger effects over individuals' decisions to self-sacrifice for the group (Swann, Gómez, et al., 2014).

Further testing of the effects of identity fusion with samples of university students across 11 countries revealed robust relationships between identity fusion with one's country and willingness to fight and die for his or her country (Swann, Buhrmester, et al., 2014). This relationship holds also during armed conflicts and warfare. Whitehouse et al. (2014) surveyed revolutionaries during the 2011 conflict in Libya and found that fusion with one's battalion predicted willingness to engage in extreme behaviors on behalf of the battalion. This effect was even stronger for revolutionaries who were more fused with their battalion than with their families, suggesting that identity fusion may capitalize on fictive-kinship psychology. Indeed, increasing the salience of genetic relatedness or psychological similarity to other countrymen multiplied the effects of identity fusion on extreme pro-group behavior (Swann, Buhrmester, et al., 2014), a finding that was further supported by Vazquez et al. (Vázquez, Gómez, Ordoñana, Swann, & Whitehouse, 2017) who showed that monozygotic twins are more fused and were more willing to make sacrifices for the other twin than dizygotic twins.

Summarizing these empirical finings, Whitehouse (2018) recently proposed a theoretical model that aims to delineate the causal chain of events that leads to identity fusion with a group and, in turn, leads to self-sacrifice for that group. At the beginning of this model are dysphoric experiences shared with other members of the group (e.g., extreme initiation rites or war-related distress), which trigger exegetical reflection on the meaning of shared suffering and transform personal identity such that people perceive the dysphoric experiences as the essence of shared group membership (Jong, Whitehouse, Kavanagh, & Lane, 2015; Whitehouse et al., 2017). The experiential component of group membership causes identity fusion with the group, and fusion translates into psychological kinship with group members, which--interacting with an outgroup threat--may lead to self-sacrifice. In other words, fused people act as if other members of their group would be their genetic relatives--willing to sacrifice one's life, especially in the context of inter-group warfare.

Here, we assess the last part of Whitehouse's model, focusing on the role of outgroup relationships in moderating the effects of fusion with one's group on sacrificial behavior. We capitalize on the intuitive appeal of the pictorial fusion measure (Fig. S1) that allowed us to deploy fusion measurements across a host of small-scale societies, ranging from hunter-gathers, over pastoralists, to market-integrated (see Table 1). As the verbal fusion measure would be problematic at sites that do not use numeric representations (hence, answering on a scale is counterintuitive), the pictorial measure allowed us to assess fusion with ingroup and outgroup across all eight societies.

However, instead of testing willingness to fight and die for one's ingroup as is common in the identity fusion literature, we operationalized self-sacrificial behavior in economic terms as coins that one is willing to give up for other members of their ethnic and religious group. That is, we employed the Random Allocation Game (RAG henceforth; Hruschka et al., 2014; Purzycki et al., 2016b) where participants allocated 30 coins between themselves and anonymous members of their religious group residing in a geographically distant place (i.e., unlikely to ever reciprocate or retaliate). Utilizing natural dynamics between selfish and pro-group behavior, this conceptualization of self-sacrifice as coins one sacrificed to the religious ingroup allowed us to measure continuous sacrificial behavior that should scale on the continuous pictorial measure of identity fusion. This avoids the assumption that self-sacrifice is a trait-like characteristic that needs a special, trait-like explanation (as assumed in Whitehouse's model). Since the endowment presented a strong incentive for participants (30 coins were roughly equivalent to half a day's wage of the local average income, except for the Hadza, who used tokens each worth 8 oz. of maize), we were able to overcome problems and biases inherent in self-reports (e.g., see Lang, Bahna, Shaver, Reddish, & Xygalatas, 2017) and measure actual self-sacrificial behavior.

Furthermore, rather than treating outgroup threat as a group-level exogenous variable imposed on participants, we measured individual-level relationships to religious outgroup using the same pictorial measure as for assessing ingroup fusion, making these measures directly comparable (see Table 1 for raw fusion means across our sites and Fig. S2 for density plots). This is not to say that the group-level variation in outgroup threat is unimportant, quite the contrary; but rather than imposing an outgroup threat artificially at each site or specifically targeting samples afflicted with intergroup aggression, we let it emerge from individual-level measures by employing varying effects of outgroup relations across our sites. The same is true for our ingroup fusion measure, allowing us to assess whether the hypothesized Ingroup Fusion*Outgroup Fusion interaction will emerge at the individual level after accounting for site-specific variance of this relationship (see section S3.1). That is, we assessed whether increasing ingroup fusion together with decreasing outgroup fusion will predict higher rates of coin sacrifice that benefits religious and co-ethnic ingroups.

2. Methods

2.1. Data availability

The data we use are from the publicly available Evolution of Religion and Morality data set (Purzycki et al., 2016a). This article, the data set, all protocol materials, site descriptions, and sampling procedures are available here: https://github.com/bgpurzycki/Evolution-of-Religion-and-Morality. The workflow for this specific project is and will be maintained at https://github.com/bgpurzycki/fusion. For results using this dataset for other purposes with other controls, see (Purzycki et al., 2016b, 2018).

2.2. Measures

2.2.1. The Random Allocation Game and sacrifice

Our behavioral measure was the Random Allocation Game, which detects systematic dishonest favoritism (Cohn, Fehr, & Maréchal, 2014; Hruschka et al., 2014; Jiang, 2013). As described in the main text, participants have the opportunity to put more coins into their own cups than chance would allow. Coins should follow a binomial distribution, but participants tend to systematically favor themselves. As such, foregoing this opportunity is a sacrifice of real gains. Moreover, the recipients are distant ingroup members; co-ethnic, co-religious individuals who share the same beliefs and backgrounds as participants, but they are not likely to directly reciprocate or retaliate. Rather than make sacrifices for ingroup members who may reciprocate at a later time, the sacrifices we measure are completely lost opportunities of gaining wealth. Participants played two games: the Self Game (focused on here) and the Local Ingroup Game, where participants allocated coins between another anonymous distant ingroup member and an anonymous local ingroup member.

2.2.2. Intergroup relations

To measure fusion, we used the visual scale from Swann Jr. et al. (2009; figure S1) ported to numeric values of 1 to 5. The questions we asked with this scale were as follows (variable names in data set in parentheses):

- Using these pictures, how emotionally close do you feel to ____________ [a specifically defined geographically distant co-ethnic, co-religionist group?
- Using these pictures, how emotionally close do you feel toward members of __________ ______ [local co-religious, co-ethnic] group?
- Using these pictures, how emotionally close do you feel toward ______________ [specifically defined non-local religious and ethnic outgroup members]?

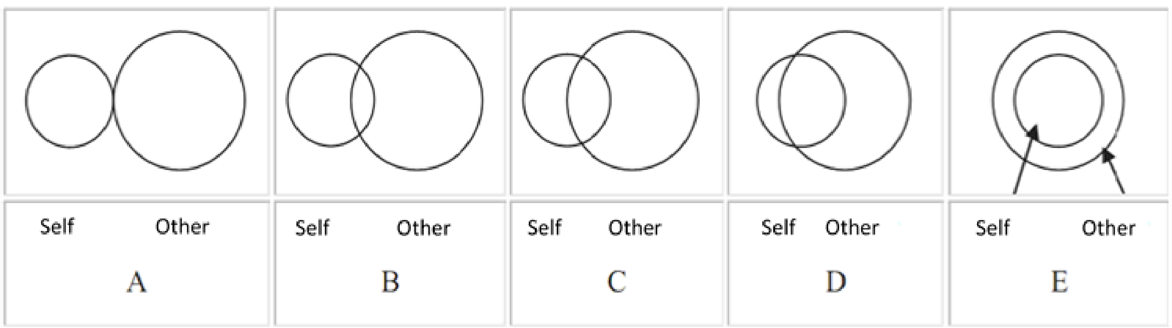

Figure S1 | Visual fusion scale adapted from Swann, et al. (2009) modified from Schubert and Otten (2002). Participants pointed to that which best represents their attitudes towards various others.

We also asked a question about participants' ratings of religious similarity to distant recipients (scale from -2 to 2):

- How similar are DISTANT's traditions/religious beliefs and practices with the LOCAL? (CORELSIM)

We do note that we are unsure as to whether or not overlapping circles function adequately as a metaphor for social relations across all contexts. In other words, the fusion measure might be too heavily reliant on a non-universal spatial metaphor. As illustrated below, we nevertheless found that by and large, individuals had far lower scores for ingroup fusion than its outgroup counterpart.

2.2.3. Local vs. extended fusion

The theory of identity fusion sketched above predicts that identity fusion is effective in facilitating pro-group behavior in interactions with both the local and extended ingroups (Swann, Buhrmester, et al., 2014; Swann, Gómez, Dovidio, et al., 2010). While participants are fused mainly with their ingroup (e.g. other Spaniards in the case of Swann et al. studies), they are often willing to sacrifice themselves even for members of groups with whom they share superordinate group membership (Europeans). That is, the pro-kin bias apparent in interaction with local ingroups (Hamilton, 1964) is projected onto members of the extended ingroup, especially if they share common cultural or morphological characteristics (Swann, Buhrmester, et al., 2014). This notion has been further elaborated in Whitehouse's model (Whitehouse, 2018), which posits that participants must first experience local fusion evoked by shared dysphoric experiences (within their community) that can be later projected onto extended ingroups.

In our current setup of the RAG, the money recipients from distant communities are extended ingroups, insofar as they are co-ethnic and co-religionist. We purposely selected distant ingroups due to having non-reciprocal relations participants because allocating to them lends itself to sacrifice inasmuch as the costs will never be directly reciprocated by those benefitting from them and it controls for any outgroup confounds. On the contrary, sacrificing resources to one's ingroup members typically involves returns through nonkin (Trivers, 1971) and kin alike (Hamilton, 1964). By way of analogy, engaging in warfare can benefit a multitude of anonymous others—who aren't fighting—from an external threat with no direct opportunity to reciprocate. Making a sacrifice of resources to one's community however, can have direct benefits, thus giving the allocations an investment-like quality instead of sacrifice.

Furthermore, we focused on ingroup fusion here primarily for the fact that the theory explicitly predicts that local fusion will predict sacrifice and represents the crucial basis of extended fusion. We nevertheless assessed the role of distant fusion on allocations by running the same main model and the model with only a simple and varying effect of local ingroup fusion (m1a and mg above) but replacing local ingroup fusion with distant ingroup fusion. Table S3 presents those results (section 3.4.2).

2.3. Data notes

While we have data for the Hadza's fusion scores (as it was a visual scale to which participants pointed), we do not have data for their perceived religious similarity due to difficulty with numerical scales. As indicated by Figure S2, the distributions of their fusion scores were intuitive insofar as they felt emotionally closer to other Hadza than to Datoga herders. In the original data file, their scale was from 0 to 4, which was recoded to 1 to 5 to be consistent with the rest of the data. One individual in the Tyvan sample responded at the halfway point between 4 and 5 on the ingroup fusion scale. This was recoded as 5.

2.4. Descriptive reporting

Figure S2 is a density plot showing the cross-site distribution of ingroup and outgroup fusion scores. Across sites, it is clear that people are rating their emotional proximity to ingroups more than their outgroups. In other words, participants were systematically more fused with their ingroups than outgroups (see main text for statistics).

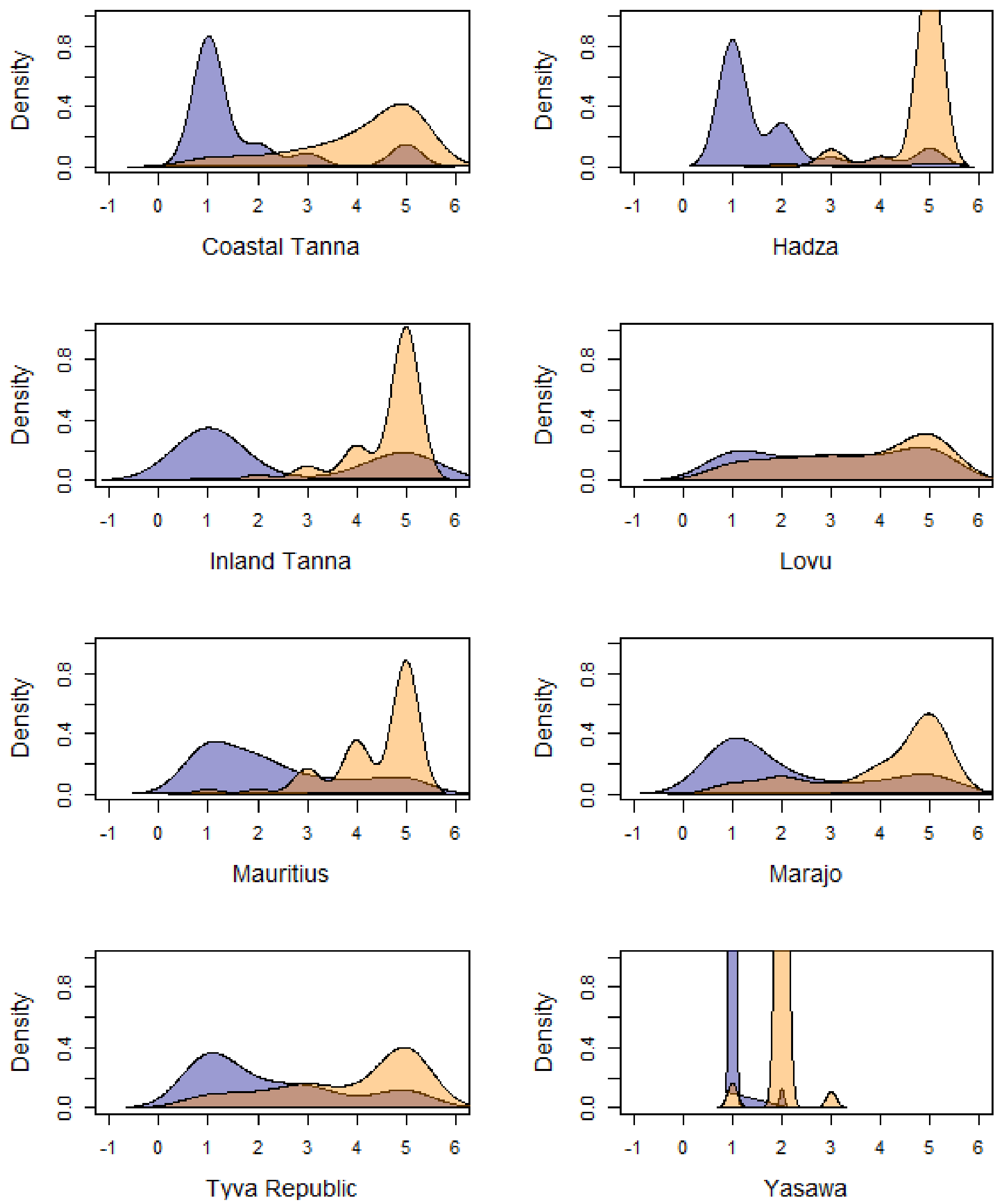

Figure S2 | Density plots of ingroup (orange) and outgroup (blue) fusion across eight field sites. Recoded scales were from 1 to 5.

Table S1 is a table of counts for participant Ingroup*Outgroup fusion scores. The modal response in the matrix is high ingroup fusion (5) and low outgroup fusion (1), totaling 146 (26% of the sample). Figure S3 includes 2D surface plot of the raw allocation data across values for ingroup and outgroup fusion. Note that the data were interpolated in order to show continuous transitions between the outgroup and ingroup categories. Note too, that the peaks are misleading insofar as the plot surfaces are means. The two peaks represent a total of three participants (e.g., one individual with low ingroup fusion and high outgroup fusion put 20 coins in the cup opposite to him or herself); see Table S1.

Table S1 | Participant responses to fusion questions of ingroups and outgroups. Left panel reports counts of participants and right panel reports means. 1 = lowest possible score (separate), 5 = maximum (fused). Bold values indicate the values of the two peaks in the plots.

	Ingroup Fusion						**Ingroup Fusion**					
Outgroup Fusion	1	2	3	4	5	*n*	1	2	3	4	5	*M*
5	**1**	4	4	3	20	*32*	**20.00**	14.50	16.00	14.33	13.05	13.94
4	1	2	**2**	6	92	*103*	16.00	16.00	**19.00**	12.17	12.48	12.69
3	2	5	12	13	27	*59*	16.50	13.60	15.25	12.77	14.00	14.03
2	6	12	12	16	32	*78*	14.33	13.33	14.17	14.94	14.09	14.18
1	81	12	20	40	146	*299*	14.78	15.08	13.80	13.60	13.85	14.11
n	91	35	50	78	317	571	14.86	14.26	14.62	13.65	13.44	13.85

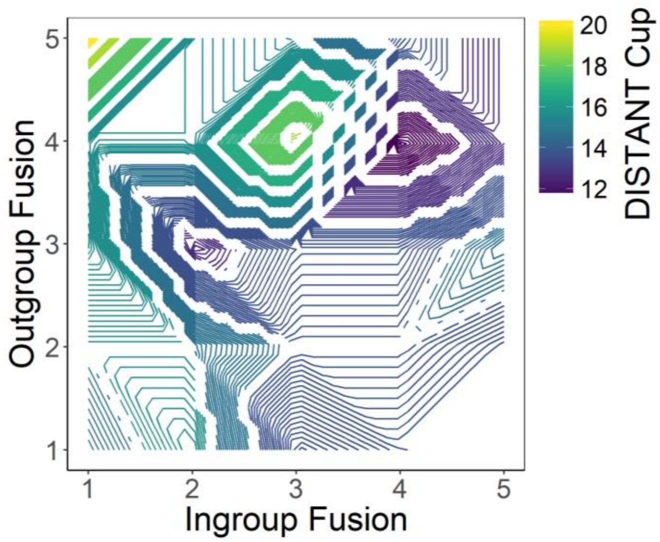

Figure S3 | 2D surface plot of raw data showing the interaction between ingroup and outgroup fusion in predicting DISTANT allocations. See Table S1 for data.

3. Main Analyses

3.1. Model definitions

Recall that the linear component of our main model was defined as follows:

$$\text{logit}(p_i) = \alpha_{S(i)} + \beta_{S(i)}*g_i + \gamma_{S(i)}*o_i + \psi_{S(i)}*g_i*o_i \qquad \text{(m1a)}$$

where g is the ingroup fusion value and o is the outgroup counterpart. We modelled their effects simply for individuals, denoted by subscript i, but also varying across groups, as denoted by the subscript s. The three other model specifications were as follows:

$$\text{logit}(p_i) = \alpha_{S(i)} + \lambda_s*r_i + \beta_{S(i)}*g_i + \gamma_{S(i)}*o_i + \psi_{S(i)}*g_i*o_i \qquad \text{(m1b)}$$

$$\text{logit}(p_i) = \alpha_{S(i)} + \beta_{S(i)}*g_i \qquad \text{(mg)}$$

$$\text{logit}(p_i) = \alpha_{S(i)} + \gamma_{S(i)}*o_i \qquad \text{(mo)}$$

Equation m1b is the same as equation m1a, but adds the variable *r*, which represents reported religious similarity of distant players. Equations mg and mo are the simplest models, where ingroup and outgroup fusion are treated as simple effects and effects varying across sites, respectively. The prior distribution definitions remain the same. We use the LKJCorr prior (where $\eta = 4$) to conservatively address extreme correlations (Lewandowski, Kurowicka, & Joe, 2009; see McElreath, 2018, pp. 393-394).

3.2. Analytical notes

The analytical strategy we adopt here primarily focuses on the theories defined in the fusion literature. That is, we stick closely to the theory that has been developed by modelling it to the best of our ability, and *not* including other factors (e.g., conventional demographic "controls") that may affect the game outcomes or absorb the effects. Specifically, Whitehouse (2018) predicts that perceived sharedness leads to local fusion which in turn leads to psychological kinship. Finally, outgroup threat moderates the relationship between psychological kinship and self-sacrifice. While we were not able to directly assess the perceived psychological kinship, we measured local fusion, hence we should observe the same effects as if we would use perceived psychological kinship (given it is a mediator).

To model monotonic effects, we used the brms package (Bürkner, 2017; Version 2.1.0) for use in R (R Core Team, 2016). We attempted to model both outgroup and ingroup fusion as independent varying effects across groups, but models would not run, likely due to their inverse correlation and distributions (see Figure S2 above). All plots other than the density plots were created using the package ggplot2 (Wickham, 2009). Note that specific values in the Bayesian models may change slightly across software as the brms package and rstan set seeds differently. We set global and per-model seeds at 7.

3.3. Results

Table S2 reports the results from the main regression models. The g_i*o_i interaction is easiest to view in Figure S4, where the highest concentration of sacrificed coins—in yellow—is when outgroup fusion is high (cf. Figure S2). Holding the effect of religious similarity constant (and removing the Hadza) shifts the concentration of sacrifice toward the higher end of the ingroup fusion scale. Figure S5 then displays the same results from model m1a across three levels of outgroup fusion (min, middle, max) with 95% credibility intervals. The relatively wide credibility intervals for maximal levels of outgroup fusion indicate that the predicted high allocations in this group (cf. Figure S4) were quite variable. Overall, however, these figures indicate that there was no difference in ingroup sacrifice across the outgroup fusion measure.

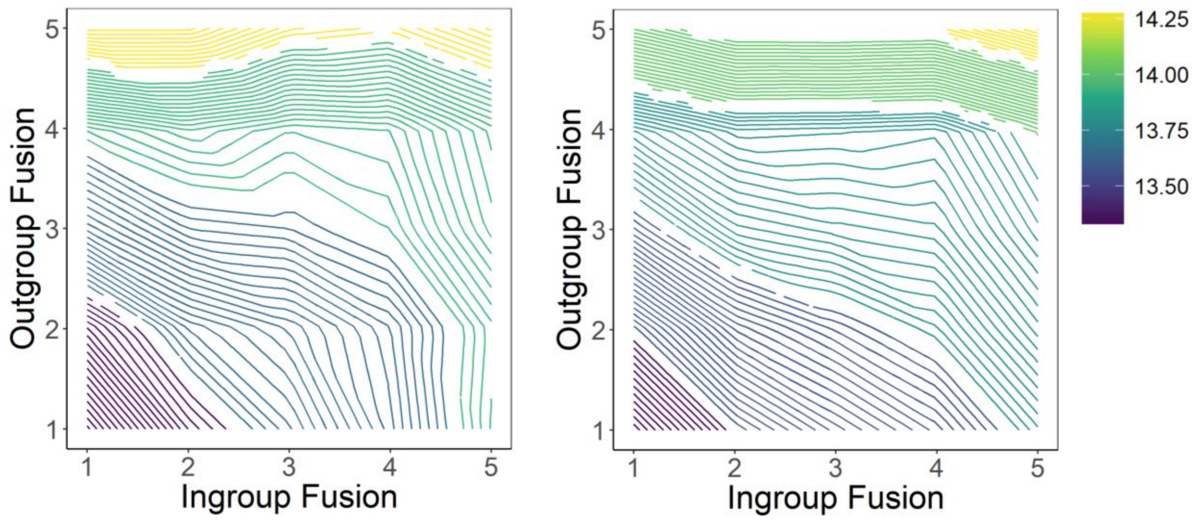

Figure S4 | Surface plot of individual-level interaction between ingroup and outgroup fusion on allocation to distant ingroup cup. X-axis is ingroup fusion (g) and y-axis is outgroup fusion (o). Left plot is model m1a and right plot is m1b.

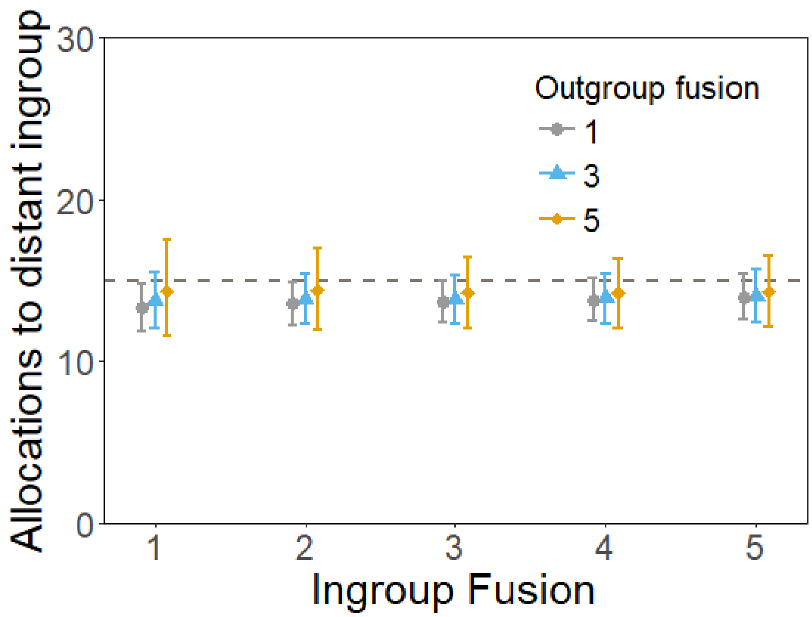

Figure S5 | Estimated means with 95% credibility intervals of allocations to distant ingroups across three levels of outgroup fusion. Estimates from model m1a. The dashed line indicates impartial allocations (half of endowment) to religious ingroups.

3.4. Supplementary models

3.4.1. Extensions to the main models

In addition to the four main models, we ran a few additional model specifications for robustness checking purposes. In all of these models, we maximize sample size by not including the religious similarity variable, which would force dropping the Hadza. The first model (model ms in Table S2) varies only outgroup fusion across sites (instead of the Outgroup*Ingroup fusion interaction), thus leaving the interaction effect only at the level of individuals:

$$\text{logit}(p_i) = \alpha_{S(i)} + \gamma_{S(i)}*o_i + \psi_i*g_i*o_i \qquad \text{(ms)}$$

We also added model specifications identical to the main model (m1a), but with either one of two experimental checks included. As all participants played at least two counterbalanced variations of the experimental game (one where the cups specified were for ingroups instead of for participants and the other cup was for another anonymous distant co-religionist, co-ethnic), we hold constant the game order (model mt in Table S2). Some participants (Lovu, Mauritius, Tyva Republic, and Yasawans) also played in a treatment condition with religious primes (model mp in Table S2). We hold this constant as well. Note that across model specifications, the main results are robust to such controls and the results are qualitatively the same across all specifications.

3.4.2. Local vs. extended fusion models

Recall in section 2.2.3 where we discussed the distinction between local fusion and extended fusion. Table S3 reports these details. The results are qualitatively similar to the results of the target models; the interaction shows no association with game outcome, but distant ingroup fusion is associated with greater sacrifice.

Table S2 | Estimates and 95% credibility intervals of models reported in main text (Figure 1). Varying effects for ingroup fusion are denoted with (g) and outgroup fusion with an (o).

	β [95% CI]	**β [95% CI]**	**β [95% CI]**	**β [95% CI]**	**β [95% CI]**	**β [95% CI]**	**β [95% CI]**
Intercept	-0.22 [-0.42, -0.03]	-0.23 [-0.44, -0.02]	-0.22 [-0.42, -0.03]	-0.17 [-0.33, -0.01]	-0.23 [-0.42, -0.05]	-0.22 [-0.43, -0.02]	-0.23 [-0.43, -0.03]
Ing.*Outg.	-0.10 [-0.52, 0.23]	-0.05 [-0.47, 0.30]	--	--	-0.09 [-0.45, 0.19]	-0.11 [-0.54, 0.23]	-0.10 [-0.51, 0.24]
Rel. Similarity	--	0.07 [-0.07, 0.20]	--	--	--	--	--
Ingroup Fusion	0.08 [-0.09, 0.25]	0.07 [-0.12, 0.27]	0.09 [-0.07, 0.25]	--	0.08 [-0.04, 0.23]	0.08 [-0.09, 0.27]	0.08 [-0.09, 0.26]
Outgroup Fusion	0.14 [-0.20, 0.56]	0.14 [-0.20, 0.54]	--	0.07 [-0.12, 0.24]	0.14 [-0.19, 0.52]	0.15 [-0.20, 0.59]	0.14 [-0.21, 0.54]
Treatment	--	--	--	--	--	-0.02 [-0.09, 0.06]	--
Game Order	--	--	--	--	--	--	0.00 [-0.06, 0.07]
Coastal Tannese	0.14 [-0.08, 0.40]	0.11 [-0.11, 0.39]	0.13 [-0.09, 0.43]	0.12 [-0.06, 0.31]	0.12 [-0.07, 0.31]	0.14 [-0.08, 0.42]	0.14 [-0.08, 0.42]
Hadza	-0.19 [-0.50, 0.06]	--	-0.22 [-0.52, 0.04]	-0.18 [-0.37, -0.00]	-0.19 [-0.38, -0.01]	-0.19 [-0.50, 0.07]	-0.18 [-0.49, 0.08]
Inland Tannese	-0.03 [-0.28, 0.20]	-0.06 [-0.31, 0.17]	-0.02 [-0.29, 0.22]	0.01 [-0.17, 0.18]	-0.00 [-0.18, 0.18]	-0.03 [-0.30, 0.22]	-0.02 [-0.28, 0.21]
Indo-Fijians	0.11 [-0.11, 0.32]	0.08 [-0.13, 0.29]	0.12 [-0.09, 0.33]	0.12 [-0.06, 0.30]	0.13 [-0.05, 0.32]	0.11 [-0.12, 0.33]	0.11 [-0.11, 0.33]
Marajó Brazilians	0.04 [-0.20, 0.25]	0.02 [-0.24, 0.23]	0.06 [-0.20, 0.28]	0.07 [-0.11, 0.25]	0.07 [-0.11, 0.26]	0.04 [-0.22, 0.26]	0.04 [-0.20, 0.25]
Mauritians	-0.00 [-0.22, 0.27]	-0.02 [-0.24, 0.26]	0.03 [-0.19, 0.32]	-0.05 [-0.24, 0.12]	-0.06 [-0.24, 0.12]	0.01 [-0.23, 0.30]	0.00 [-0.23, 0.28]
Tyvans	0.16 [-0.04, 0.39]	0.11 [-0.09, 0.34]	0.19 [-0.01, 0.42]	0.17 [-0.00, 0.35]	0.17 [-0.00, 0.35]	0.17 [-0.05, 0.39]	0.17 [-0.05, 0.39]
Yasawan-Fijians	-0.26 [-0.50, -0.05]	-0.25 [-0.52, -0.04]	-0.27 [-0.52, -0.07]	-0.28 [-0.46, -0.11]	-0.24 [-0.43, -0.07]	-0.26 [-0.52, -0.05]	-0.25 [-0.51, -0.05]
Coastal Tannese[v]	-0.03 [-0.37, 0.24]	-0.05 [-0.39, 0.19]	-0.05 [-0.36, 0.12]	-0.11 [-0.45, 0.12]	-0.10 [-0.44, 0.12]	-0.04 [-0.38, 0.24]	-0.04 [-0.40, 0.25]
Hadza[v]	-0.06 [-0.46, 0.25]	--	-0.01 [-0.27, 0.25]	-0.18 [-0.56, 0.04]	-0.16 [-0.53, 0.05]	-0.07 [-0.48, 0.25]	-0.06 [-0.46, 0.27]
Inland Tannese[v]	0.03 [-0.23, 0.35]	0.02 [-0.23, 0.30]	0.04 [-0.16, 0.30]	0.03 [-0.16, 0.26]	0.04 [-0.17, 0.28]	0.03 [-0.26, 0.37]	0.02 [-0.26, 0.34]
Indo-Fijians[v]	0.00 [-0.26, 0.26]	-0.01 [-0.27, 0.23]	0.03 [-0.13, 0.25]	0.01 [-0.22, 0.24]	-0.00 [-0.23, 0.22]	0.00 [-0.26, 0.27]	0.00 [-0.25, 0.26]
Marajó Brazilians[v]	0.01 [-0.27, 0.29]	-0.00 [-0.28, 0.25]	0.05 [-0.12, 0.30]	0.08 [-0.13, 0.35]	0.07 [-0.14, 0.34]	0.01 [-0.26, 0.30]	0.01 [-0.27, 0.30]
Mauritians[v]	0.04 [-0.20, 0.35]	0.04 [-0.20, 0.31]	-0.07 [-0.37, 0.09]	0.10 [-0.09, 0.38]	0.11 [-0.09, 0.40]	0.05 [-0.22, 0.37]	0.04 [-0.22, 0.36]
Tyvans[v]	-0.00 [-0.28, 0.26]	-0.00 [-0.24, 0.24]	-0.01 [-0.21, 0.17]	0.04 [-0.18, 0.29]	0.04 [-0.19, 0.30]	-0.00 [-0.29, 0.26]	-0.01 [-0.29, 0.25]
Yasawa-Fijians[v]	0.01 [-0.39, 0.42]	0.02 [-0.34, 0.43]	0.02 [-0.24, 0.35]	0.03 [-0.41, 0.56]	0.02 [-0.40, 0.53]	0.01 [-0.40, 0.44]	0.01 [-0.37, 0.42]
[v]Varied effect	g*o	g*o	g	o	o	g*o	g*o
Panel from Fig. 1	a	b	c	d	--	--	--
Model (R script)	m1a	m1b	mg	mo	ms	mp	mt

Table S3| Estimates and 95% credibility intervals of estimates for distant ingroup fusion models. Varying effects for distant ingroup fusion are denoted with (c) and outgroup fusion with an (o).

	β [95% CI]	**β [95% CI]**
Intercept	-0.19 [-0.37, -0.02]	-0.19 [-0.38, -0.00]
Dist.*Outg.	-0.06 [-0.40, 0.24]	--
Distant Fusion	0.04 [-0.09, 0.17]	0.06 [-0.05, 0.17]
Outgroup Fusion	0.11 [-0.17, 0.38]	--
Treatment	--	--
Game Order	--	--
Coastal Tannese	0.14 [-0.06, 0.37]	0.11 [-0.10, 0.35]
Hadza	-0.18 [-0.41, 0.02]	-0.23 [-0.46, -0.01]
Inland Tannese	0.00 [-0.20, 0.20]	0.01 [-0.21, 0.21]
Indo-Fijians	0.12 [-0.07, 0.32]	0.14 [-0.06, 0.34]
Marajó Brazilians	0.07 [-0.12, 0.26]	0.10 [-0.11, 0.30]
Mauritians	-0.06 [-0.25, 0.13]	-0.03 [-0.24, 0.16]
Tyvans	0.18 [-0.01, 0.37]	0.19 [-0.01, 0.39]
Yasawan-Fijians	-0.28 [-0.52, -0.08]	-0.30 [-0.55, -0.09]
Coastal Tannese^{v}	-0.06 [-0.45, 0.25]	-0.04 [-0.24, 0.07]
Hadza^{v}	-0.14 [-0.60, 0.17]	-0.00 [-0.14, 0.15]
Inland Tannese^{v}	-0.01 [-0.43, 0.31]	0.01 [-0.11, 0.15]
Indo-Fijians^{v}	0.02 [-0.25, 0.31]	-0.00 [-0.14, 0.12]
Marajó Brazilians^{v}	0.09 [-0.15, 0.47]	0.02 [-0.10, 0.18]
Mauritians^{v}	0.08 [-0.18, 0.42]	0.02 [-0.09, 0.16]
Tyvans^{v}	0.01 [-0.29, 0.31]	0.00 [-0.12, 0.13]
Yasawa-Fijians^{v}	0.03 [-0.44, 0.61]	-0.00 [-0.18, 0.19]
^{v}Varied effect	c*o	c
Model (R script)	mc1	mc2

3.4.3. The Individualist Game in the Tyva Republic

In the main text, we speculated that the Ingroup*Outgroup interaction may be clearer in a self vs. ingroup game dyad. In this dyad, rather than a geographically distant ingroup, participants would sacrifice coins to their *local* ingroup at a cost to themselves. As we noted, Tyvans (n = 81) played this game, dubbed the "Individualist Game" (Purzycki & Kulundary, 2018). However, the fusion scores were not the focal component of the initial report. Here, we apply a version of the main model reported in the text to the Tyvan data (m1tyva) along with a version without the proposed interaction (m2tyva):

$$y_i \sim \text{Binomial}(30, p_i)$$

$$\text{logit}(p_i) = \alpha + \beta_g * g_i + \beta_o * o_i + \beta_{go} g_i * o_i \quad \text{(m1tyva)}$$

$$\text{logit}(p_i) = \alpha + \beta_g * g_i + \beta_o * o_i \quad \text{(m2tyva)}$$

$$\alpha, \beta_g * g_i, \beta_o * o_i, \beta_{go} g_i * o_i \sim \text{Normal}(0,1)$$

Here, y_i denotes the allocations to Buddhist Tyvans in Kyzyl, g denotes fusion with Buddhist Tyvans from Kyzyl, and o denotes fusion with ethnic Russians from Ak Dovurak, a mining town in western Tyva. Again, we modelled effects monotonically. As this was a single site, we varied neither intercepts nor effects.

Table S4 | Estimates and 95% credibility intervals of model defined above with sample from the Tyva Republic. Model names correspond to the code in the supplementary R script.

	m1tyva	**m2tyva**
Intercept	0.04 [-0.51, 0.37]	0.09 [-0.28, 0.31]
Outgroup	0.02 [-0.61, 0.79]	0.05 [-0.26, 0.53]
Ingroup	-0.10 [-0.64, 0.57]	-0.13 [-0.35, 0.12]
Outgroup * Ingroup	0.04 [-0.84, 0.91]	--

Table S4 reports the output showing that the ingroup-outgroup fusion interaction had no association with outcome in the games (m1tyva). Note, however, that ingroup fusion did predict *withholding* coins from the local group. Note, too, that Tyvans largely played fairly (if a little generous; mean allocations to the ingroup was 15.23, SD = 2.88), thus making it difficult to infer what similar effects would be like in contexts with greater general withholding. Other analyses of the Tyvan data are reported elsewhere, but the fusion measures were not a focus of those analyses (Purzycki & Kulundary, 2018).

4. Supplementary Frequentist Analyses

In this section, we provide additional analyses using frequentist analytical techniques as a robustness check for the findings produced by our main Bayesian models. First, as a quasimanipulation-check, we examined the correlation between ingroup and outgroup fusion, assessing whether heightened ingroup fusion comes at the expense of decreased outgroup fusion. Contrary to this prediction, we observed a moderate positive correlation between these two measures (Pearson's r = 0.28, 95% CI = [0.21 – 0.35]), suggesting that the pictorial measure may capture general sociality rather than relationship to specific groups. Despite this finding, 26% of our sample indicated maximum fusion with ingroup and simultaneously minimal fusion with outgroup whereas only 14% indicated general prosociality (maximum on both fusion measures, see Tab. S1). This result indicates that a sufficiently high number of participants reported the purported negative relationship, lending credibility to our subsequent analyses.

First, we analyzed the simple effects of ingroup and outgroup fusion on allocations to distant ingroups, varying these effects across sites. These analyses revealed that an increase of one on the pictorial ingroup fusion measure was associated with an increase of 0.6% in the likelihood of allocating a coin to religious ingroups (instead to the self). That is, the increase from minimum to maximum levels of ingroup fusion was associated with a 3% increase in the likelihood of coin sacrifice. Similar results were obtained for the outgroup fusion measure, where an increase of one was associated with 0.4% increase in the likelihood of allocating to distant ingroups. In addition to the aforementioned correlation, these results also suggest that, at least in our sites, the measure of fusion was not sensitive to targeted groups (ingroup vs. outgroup), and instead measured general prosociality. Indeed, the coefficients of both measures were positive (see Table S5, Models fg and fo).

However, it could be argued that our data were generated by two processes, which mask the purported effects of outgroup hostility. That is, the effects of outgroup hostility on extreme ingroup sacrifice can be observed only for participants reporting high levels of ingroup fusion and low levels of outgroup fusion; whereas reporting a mixture of positive ingroup and outgroup fusion leads to a small sacrifice corresponding to general prosociality. To investigate this proposition, we interacted ingroup and outgroup fusion and assessed the effects of this interaction on coin allocations to religious ingroups. The raw data (ignoring within-site nesting) are displayed in Fig. S3. The surface plots reveal that the highest allocations were in the segment of maximal fusion with outgroups and minimal fusion with ingroups. On the contrary, the predicted region of interest (high ingroup and low outgroup fusion) exhibited low mean allocations. However, recall that the raw data are misleading due to low number of participants in some of these extreme segments (see Tab. S1).

To investigate this matter more rigorously, we attempted to build a multi-level binomial model, varying the Ingroup*Outgroup interaction across sites. However, fitting this model in R (R Core Team, 2016) using the *glmer* function in the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015, p. 4) revealed poor convergence of this model, despite trying several optimizers. Hence, we were forced to divide our outgroup fusion measure into simpler categories (1-2 = 0; 3-4 = 1; 5 = 2) in order to reach convergence. This problem bolsters the utility of the Bayesian framework when analyzing multi-level data with effects varying across nesting factors. Also note that compared to the Bayesian models m1a and m1b, we treat the ingroup and outgroup fusion measures as continuous in the frequentist analyses (instead of modelling them as monotonic effects).

The results of the interaction model suggest that the slope of ingroup fusion predicting allocations to religious ingroups is more positive for low levels of outgroup fusion (see Table S5). However, as can be seen in Fig. S6, these differences are driven by participants reporting low ingroup fusion; the intercepts differ across the three levels of outgroup fusion such that higher outgroup fusion is associated with higher allocations to distant ethnic and religious ingroups (cf. Figure S5). In other words, for maximum levels of ingroup fusion, outgroup fusion does not play a role because all participants tended to split the endowment impartially between themselves and distant ingroups (while the theory predicts highest allocations for low levels of outgroup fusion). Instead, the difference in slopes was driven by participants who indicated low levels of both ingroup and outgroup fusion; such participants were the most selfish. For completeness, we also report the same analyses using fusion with distant ingroups rather than local ingroups in Table S5 (cf. Table S3).

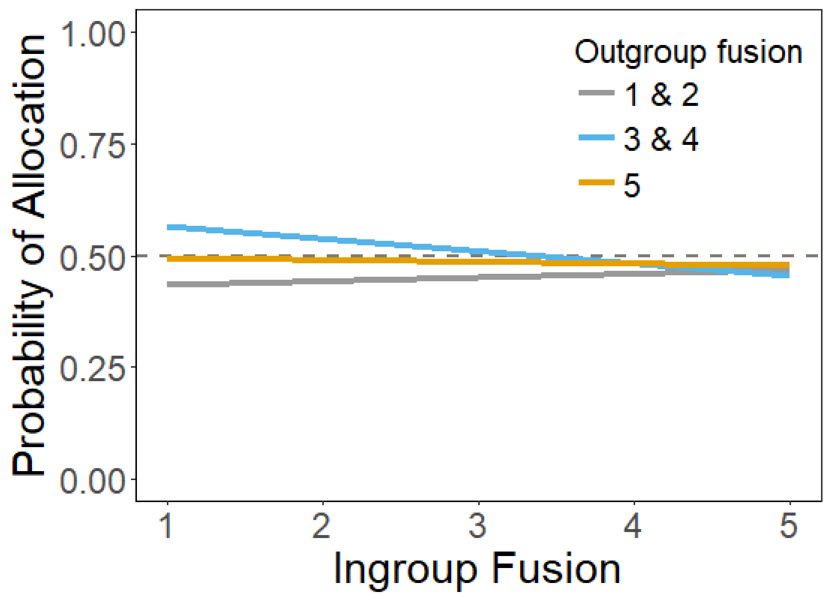

Figure S6 | Estimated regression lines from the multi-level binomial model across three levels of outgroup fusion. While the slope of ingroup fusion when outgroup fusion equals one or two is more positive compared to slopes across other levels of outgroup fusion, this effect does not lead to larger amounts of sacrificed coins. Quite the contrary, maximal ingroup fusion led to impartial coin allocations across all levels of outgroup fusion, indicating that the pictorial fusion measure assesses general prosociality. It is the lack of fusion with any group that drives the difference – such participants are more selfish. Note: the y-axis displays logistic transformations of the raw estimates to indicate probability of allocating a coin to religious ingroup. The dashed line indicates impartial allocations (half of endowment) to religious ingroups.

Table S5 | Estimates and 95% confidence intervals for frequentist models. Varying effects for ingroup fusion are denoted with (g) and outgroup fusion with an (o). We also present models with distant fusion (c).

	Ingroup models			**Distant models**	
	β [95% CI]	**β [95% CI]**	**β [95% CI]**	**β [95% CI]**	**β [95% CI]**
Intercept	-0.25 [-0.44, -0.05]	-0.20 [-0.31, -0.08]	-0.30 [-0.49, -0.11]	-0.21 [-0.35, -0.06]	-0.27 [-0.50, -0.04]
Ingroup/Distant	0.03 [-0.01, 0.06]	-- --	0.04 [-0.001, 0.07]	0.02 [-0.01, 0.04]	0.03 [-0.01, 0.06]
Outgroup	-- --	0.02 [-0.01, 0.05]	-- --	-- --	-- --
Outg. = 1	-- --	-- --	0.67 [-0.04, 1.39]	-- --	0.33 [0.20., 0.85]
Outg. = 2	-- --	-- --	0.30 [-0.31, 0.92]	-- --	0.28 [-0.14, 0.69]
Ing./Dist.*Outg. = 1	-- --	-- --	-0.15 [-0.30, 0.01]	-- --	-0.06 [-0.19, 0.06]
Ing./Dist.*Outg. = 2	-- --	-- --	-0.05 [-0.18, 0.08]	-- --	-0.05 [-0.15, 0.05]
Coastal Tannese	-0.11	-0.08	-0.16	-0.12	-0.06
Hadza	-0.61	-0.35	-0.63	-0.43	-0.68
Inland Tannese	-0.24	-0.18	-0.30	-0.19	-0.24
Indo-Fijians	-0.06	-0.09	-0.15	-0.07	-0.06
Marajó Brazilians	-0.11	-0.14	-0.22	-0.09	-0.10
Mauritians	-0.29	-0.24	-0.34	-0.23	-0.29
Tyvans	0.01	-0.03	-0.07	-0.02	0.03
Yasawan-Fijians	-0.54	-0.45	-0.55	-0.50	-0.78
Coastal Tannese[v]	0.01	0.01	0.51/0.12	--	0.08/0.02
Hadza[v]	0.05	-0.01	1.07/0.75	--	0.81/0.78
Inland Tannese[v]	0.03	0.02	0.67/0.28	--	0.29/0.24
Indo-Fijians[v]	0.01	0.02	0.49/0.05	--	0.07/0.02
Marajó Brazilians[v]	0.01	0.03	0.57/0.21	--	0.12/0.06
Mauritians[v]	0.03	0.03	0.72/0.40	--	0.34/0.29
Tyvans[v]	0.01	0.03	0.39/-0.02	--	-0.04/-0.10
Yasawa-Fijians[v]	0.05	0.01	0.97/0.63	--	0.92/0.90
[v]Varied effect	g	o	g*o=1/g*o=2	c	c*o=1/c*o=2
Model (R script)	fg	fo	fgo	fc	fco

Note: Compared to Bayesian models (Table S2), the frequentist models do not display 95% CI for varying effects. Ingroup/Distant show coefficients for ingroup and distant fusion, respectively. The model fc does not contain site-specific intercepts due to convergence issues.

5. Supplementary Discussion

In summary, these results suggest that lack of fusion with any group is generally associated with selfishness whereas increasing fusion with either outgroup or ingroup increases willingness to sacrifice coins for distant ingroups. In other words, the effect of outgroups appears to be additive rather than multiplicative. In our sample, the measure of outgroup fusion worked as an indicator of general pro-group orientation. Nevertheless, these results are limited by two possibilities. First, it could be argued that the structure of the RAG, where impartial money allocation between oneself and religious ingroup is according to the rules of the game, capped possible extreme sacrificial behavior for the ingroup. Allocating more than half of one's endowment to religious ingroups would mean breaking the rules of the game, hence we could not observe extreme behaviors in the region of interest (maximal ingroup and minimal outgroup fusion). On the other hand, the identity fusion literature often works with extreme—but hypothetical—scenarios such as jumping from a bridge to stop a trolley with one's body; thus, breaking the rules in an anonymous game to sacrifice more coins for the ingroup perhaps should not be that difficult for highly fused participants; there is no *a priori* reason to discount subtler forms of sacrifice and the theory does not speak to this subtlety. Second, to avoid confounds, we purposefully selected religious and ethnic outgroups that are not in a direct hostile conflict with the ingroup. We were mostly successful. While the effects of outgroup hostility should be observed even in situations that lack direct between-group conflict (groups still compete), intensified inter-group conflict might limit the extreme outgroup fusion scores we found in our sample and motivate people to allocate more coins to ingroups.

6. References

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28.

Cohn, A., Fehr, E., & Maréchal, M. A. (2014). Business culture and dishonesty in the banking industry. *Nature*, *516*(7529), 86–89. https://doi.org/10.1038/nature13977

Gómez, A., Brooks, M. L., Buhrmester, M. D., Vázquez, A., Jetten, J., & Swann Jr, W. B. (2011). On the nature of identity fusion: insights into the construct and a new measure. *Journal of Personality and Social Psychology*, *100*(5), 918.

Hamilton, W. D. D. (1964). The genetical evolution of social behaviour, I & II. *Journal of Theoretical Biology*, *7*(1), 1–52. https://doi.org/10.1016/0022-5193(64)90038-4

Hruschka, D., Efferson, C., Jiang, T., Falletta-Cowden, A., Sigurdsson, S., McNamara, R., … Henrich, J. (2014). Impartial institutions, pathogen stress and the expanding social network. *Human Nature*, *25*(4), 567–579. https://doi.org/10.1007/s12110-014-9217-0

Jiang, T. (2013). Cheating in mind games: The subtlety of rules matters. *Journal of Economic Behavior & Organization*, *93*, 328–336. https://doi.org/10.1016/j.jebo.2013.04.003

Jong, J., Whitehouse, H., Kavanagh, C., & Lane, J. (2015). Shared Negative Experiences Lead to Identity Fusion via Personal Reflection. *PLoS ONE*, *10*(12), 1–12. https://doi.org/10.1371/journal.pone.0145611

Lang, M., Bahna, V., Shaver, J. H., Reddish, P., & Xygalatas, D. (2017). Sync to link: Endorphin-mediated synchrony effects on cooperation. *Biological Psychology*, *127*, 191–197. https://doi.org/https://doi.org/10.1016/j.biopsycho.2017.06.001

Lewandowski, D., Kurowicka, D., & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, *100*(9), 1989–2001. https://doi.org/10.1016/j.jmva.2009.04.008

McElreath, R. (2018). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. CRC Press.

Purzycki, B. G., Apicella, C., Atkinson, Q. D., Cohen, E., McNamara, R. A., Willard, A. K., … Henrich, J. (2016a). Cross-cultural dataset for the evolution of religion and morality project. *Scientific Data*, *3*, 160099. https://doi.org/10.1038/sdata.2016.99

Purzycki, B. G., Apicella, C., Atkinson, Q. D., Cohen, E., McNamara, R. A., Willard, A. K., … Henrich, J. (2016b). Moralistic gods, supernatural punishment and the expansion of human sociality. *Nature*, *530*(7590), 327–330. https://doi.org/10.1038/nature16980

Purzycki, B. G., Henrich, J., Apicella, C., Atkinson, Q. D., Baimel, A., Cohen, E., … Norenzayan, A. (2018). The evolution of religion and morality: a synthesis of ethnographic and experimental evidence from eight societies. *Religion, Brain & Behavior*, *8*(2), 101–132. https://doi.org/10.1080/2153599X.2016.1267027

Purzycki, B. G., & Kulundary, V. (2018). Buddhism, identity, and class: fairness and favoritism in the Tyva Republic. *Religion, Brain & Behavior*, *8*(2), 205–226. https://doi.org/10.1080/2153599X.2016.1267031

R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Schubert, T. W., & Otten, S. (2002). Overlap of Self, Ingroup, and Outgroup: Pictorial Measures of Self-Categorization. *Self and Identity*, *1*(4), 353–376. https://doi.org/10.1080/152988602760328012

Swann Jr., W. B., Gómez, Á., Conor, D., Francisco, J., & Huici, C. (2009). Identity fusion: The interplay of personal and social identities in extreme group behavior. *Journal of Personality and Social Psychology*, *96*(5), 995–1011. https://doi.org/10.1037/a0013668

Swann, W. B. (1983). Self-verification: Bringing social reality into harmony with the self. In *Social Psychological Perspectives on the Self, Vol. 2* (pp. 33–66). Hilsdale, NJ: Erlbaum. https://doi.org/10.1126/science.218.4574.782

Swann, W. B., Buhrmester, M. D., Gómez, A., Jetten, J., Bastian, B., Vázquez, A., … Zhang, A. (2014). What makes a group worth dying for? Identity fusion fosters perception of familial ties, promoting self-sacrifice. *Journal of Personality and Social Psychology*, *106*(6), 912–926. https://doi.org/10.1037/a0036089

Swann, W. B., Gómez, Á., Buhrmester, M. D., López-Rodríguez, L., Jiménez, J., & Vázquez, A. (2014). Contemplating the ultimate sacrifice: Identity fusion channels pro-group affect, cognition, and moral decision making. *Journal of Personality and Social Psychology*, *106*(5), 713–727. https://doi.org/10.1037/a0035809

Swann, W. B., Gómez, A., Dovidio, J. F., Hart, S., & Jetten, J. (2010). Dying and killing for one's group: identity fusion moderates responses to intergroup versions of the trolley problem. *Psychological Science*, *21*(8), 1176–1183. https://doi.org/10.1177/0956797610376656

Swann, W. B., Gómez, Á., Huici, C., Morales, J. F., & Hixon, J. G. (2010). Identity Fusion and Self-Sacrifice: Arousal as a Catalyst of Pro-Group Fighting, Dying, and Helping Behavior. *Journal of Personality and Social Psychology*, *99*(5), 824–841. https://doi.org/10.1037/a0020014

Swann, W. B., Gómez, A., Seyle, D. C., Morales, J. F., & Huici, C. (2009). Identity fusion: the interplay of personal and social identities in extreme group behavior. *Journal of Personality and Social Psychology*, *96*(5), 995–1011. https://doi.org/10.1037/a0013668

Swann, W. B., & Hill, C. A. (1982). When our identities are mistaken: Reaffirming self-conceptions through social interaction. *Journal of Personality and Social Psychology*, *43*(1), 59–66. https://doi.org/10.1037/0022-3514.43.1.59

Swann, W. B., Jetten, J., Gómez, Á., Whitehouse, H., & Bastian, B. (2012). When group membership gets personal: A theory of identity fusion. *Psychological Review*, *119*(3), 441–456. https://doi.org/10.1037/a0028589

Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, *46*(1), 35–57. https://doi.org/10.1086/406755

Vázquez, A., Gómez, Á., Ordoñana, J. R., Swann, W. B., & Whitehouse, H. (2017). Sharing genes fosters identity fusion and altruism. *Self and Identity*, *16*(6), 684–702. https://doi.org/10.1080/15298868.2017.1296887

Whitehouse, H. (2018). Dying for the group : Towards a general theory of extreme self-sacrifice. https://doi.org/10.1017/S0140525X18000249

Whitehouse, H., Jong, J., Buhrmester, M. D., Gómez, Á., Bastian, B., Kavanagh, C. M., … Gavrilets, S. (2017). The evolution of extreme cooperation via shared dysphoric experiences. *Scientific Reports*, *7*, 44292. https://doi.org/10.1038/srep44292

Whitehouse, H., McQuinn, B., Buhrmester, M., & Swann, B. W. (2014). Brothers in arms: Libyan revolutionaries bond like family. *Proceedings of the National Academy of Sciences*, (Early edition), 10–12. https://doi.org/10.1073/pnas.1416284111

Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York.