Prosociality declines after a collective ritual: Evidence from a patron saint festival in Oaxaca, Mexico

Cameron M. Curtin[a,b], Narcedalia Vasquez Martinez[c],
Yunitza Vásquez Vásquez[d], & Joseph Henrich[b]

[a] Department of Social and Behavioral Sciences, Toulouse School of Economics
[b] Department of Human Evolutionary Biology, Harvard University
[c] Division of Mechatronic and Industrial Engineering, Tecnológico de Estudios Superiores de Ecatepec
[d] Department of Anthropology, Universidad Autónoma Metropolitana

September 22, 2026

Correspondence concerning this article should be addressed to Cameron M. Curtin, Department of Social and Behavioral Sciences, Toulouse School of Economics, 1 Esplanade de l'Université, 31000 Toulouse, France. Email: cameron.curtin@tse-fr.eu

Prosociality declines after a collective ritual: Evidence from a patron saint festival in Oaxaca, Mexico

Abstract

The cross-cultural ubiquity of costly collective rituals presents a puzzle. Researchers have proposed that collective rituals enhance group solidarity, facilitating these rituals' spread through intergroup competition. Accordingly, studies indicate that collective rituals can foster cohesion and cooperation in lab and real-world settings. We test this hypothesis in a novel, naturalistic setting: patron saint festivals (*fiestas*) in indigenous communities of Oaxaca, Mexico. In this preregistered natural experiment, we used mixed methods to study changes in cohesion, prosocial attitudes, and cooperation after a *fiesta* in one Zapotec village. Our within-subjects design included 36 survey participants, 25 of whom also completed a Public Goods Game. Contrary to prior research, we find evidence of declines in altruism and cooperation, and no change in cohesion. Nonetheless, the *fiesta* is a huge cooperative endeavor. While the sample is constrained by the community's size and this naturalistic event does not offer a fully controlled field experiment, these results offer important nuance to our understanding of collective ritual.

Keywords: ritual; cooperation; cohesion; cultural evolution; field experiment

1. Introduction

Collective rituals, such as those involving communal singing, dancing, pain, or terror, are ubiquitous across human societies (Norton, 2024; Xygalatas, 2023). In the Kalahari, !Kung hunter-gatherers traditionally danced, clapped, and sang all night during elaborate healing ceremonies (Marshall, 2009[1973]). Starting in boyhood, men in Papua New Guinea underwent a series of group initiation rituals that involved genital mutilation, adhering to strict food taboos, and head-hunting victims from rival groups (Tuzin, 1976, 2001)– an example of *rites of terror* that recur across many small-scale societies (Whitehouse, 1996). In contemporary, large-scale societies, collective ritual is common in the form of group worship, hazing rituals, and patriotic parades. Given that these rituals are often costly for participants (Henrich, 2009; Sosis & Alcorta, 2003), this ubiquity poses a puzzle. How did these costly cultural practices become so widespread?

Building on ethnographers' longstanding observations that collective rituals seem to bind groups more closely together (Durkheim, 2001[1912]), one leading hypothesis proposes that collective rituals promote group cohesion and cooperation (Hobson et al., 2018; Irons, 2001; Norenzayan et al., 2016; Sosis & Alcorta, 2003; Watson-Jones & Legare, 2016; Xygalatas, 2023). According to cultural evolutionary theory, cultural practices that enhance cooperation spread through intergroup competition, including intergroup conflict (war, raiding), selective copying (less successful groups copy more successful groups), and selective migration (people in less successful groups migrate to more successful groups) (Richerson et al., 2016). Thus, psychologically potent collective rituals may have spread as groups with rituals that fostered greater cohesion and cooperation outcompeted less cooperative, cohesive groups (Norenzayan et al., 2016).

While prominent, this is not the only hypothesis about the origin of collective ritual. Complementary hypotheses suggest that collective rituals provide opportunities to signal commitment and loyalty to the group and facilitate the identification of in-group members (Hobson et al., 2018; Watson-Jones & Legare, 2016). On the other hand, some scholars have argued that collective ritual does not serve any adaptive function. For example, Liénard & Boyer (2006) suggest instead that collective rituals were selectively transmitted over generations simply because they are cognitively sticky.

Supporting the cultural evolutionary hypothesis, lab-based experiments have found that key elements of collective ritual promote group cohesion and cooperation. Many collective rituals feature behavioral synchrony (matched rhythmic movement, such as dancing or drumming) and shared physiological arousal. In experiments with adults, behavioral synchrony and shared physiological arousal increase trust, feelings of "being on the same team", affiliation, cohesion, and cooperation (Jackson et al., 2018; Mogan et al., 2017; Wiltermuth & Heath, 2009). Similar effects have been observed in children. For example, preschoolers were more likely to spontaneously help and cooperate with a partner after participating in a ritualized group activity that involved singing and walking in step (Kirschner & Tomasello, 2010). Meanwhile, repeated participation in a ritualized group activity increases children's self-reported ingroup affiliation, affiliative ingroup displays, and outgroup monitoring (Wen et al., 2016, 2020). The early emergence of these tendencies suggests that humans may come into the world psychologically prepared to engage in ritual (Wen et al., 2016).

Taking this research "into the wild", empirical studies of real-life collective rituals have found associations between collective ritual, cohesion, and cooperation. For example, in Spain, people reported increased feelings of belonging and identity fusion– a visceral feeling of

"oneness" with the group– after participating in folkloric drumming marches (Páez et al., 2015). Similarly, self-reported measures of cohesion increased during a Hindu religious festival (Singh et al., 2020), and participation in such festivals is associated with more cohesive networks of social support (Power, 2018). Providing evidence that collective rituals may enhance prosocial behavior, Israeli kibbutz members who participated in more collective prayer rituals were more cooperative in a behavioral economics game (Ruffle & Sosis, 2007). In Mauritius, observing an extreme ritual was associated with donation of a larger proportion of a cash endowment to the local temple (Xygalatas et al., 2013).

We build on this literature by testing the hypothesis that collective ritual fosters cohesion and cooperation in a novel, naturalistic setting. Every year, indigenous communities in Oaxaca, Mexico host elaborate, multi-day festivals (*fiestas*) in honor of town patron saints. *Fiestas* involve various collective ritual activities, including dancing, music-making, and processions (Figure 1). These celebrations are costly for participants in terms of money, time, and energy (de la Fuente, 1949; Stephen, 2005). There are undoubtedly a variety of proximate reasons why people participate in *fiestas*, including religious devotion, local customs, and social norms. At a more ultimate level, however, *fiestas* may serve an adaptive function by binding communities closer together and promoting cooperation. This could be particularly important in Oaxaca. In this highly indigenous state, the community serves the primary locus of identity (Martínez, 2013). Most communities self-govern to provide local public goods and manage communally-owned lands (Curtin et al., 2024). Moreover, violent intergroup conflict– especially over land– is common (Cook, 2014; Dennis, 1987; López-Bárcenas, 2004; Yannakakis, 2008). Group cohesion and cooperation are paramount to community function. Oaxacan *fiestas* thus serve as a

Figure 1. Dancing during the fiesta. (A) Women dressed in traditional garments (*huipiles*) danced traditional dances (*jarabes*) with baskets of flowers and sparklers during the opening celebration of the fiesta. Soon, other community members joined the dancing. Over the course of about 6 hours, the community and bands processed around town, stopping frequently to dance. (B) Costumed dancers performed a traditional regional dance for about 1.5 hours. Source: Cameron M. Curtin, 2022.

compelling natural experiment for studying how collective ritual shapes ingroup-oriented prosociality.

To test the hypothesis that *fiestas* enhance group cohesion and cooperation, we compared the prosocial attitudes and behavior of community members before and after the largest annual *fiesta* in one Zapotec village. Building on prior literature, we pre-registered (https://doi.org/10.17605/OSF.IO/CW693) the hypotheses that after the *fiesta*, participants would (1) feel a greater sense of cohesion with the community, (2) report more prosocial attitudes towards community members, and (3) behave more cooperatively with community members.

This study extends current scholarship in several ways. First, we expand sampling to a novel real-world setting, contributing to ongoing efforts to assess the generalizability of established findings and to sample from populations underrepresented in the psychology literature. Second, we leverage mixed methods– including psychological surveys, interviews, and

a behavioral economics game– in a preregistered within-subjects design. Previous studies of real-world collective rituals have combined only some of these elements, such as within-subjects surveys (e.g. Páez et al., 2015; Singh et al., 2020) or between-subjects surveys and/or behavioral games (e.g. Ruffle & Sosis, 2007; Xygalatas et al., 2013). Finally, detailed ethnography adds important context for interpreting the results– CMC, NVM, & YVV spent 4 months living in the community.

2. Methods

Data for this study were collected as part of a larger research project in a Zapotec village in the Sierra Norte region of Oaxaca, Mexico. We collected data in two waves: baseline (May-June 2022) and *fiesta* (October 2022). During each wave, participants were invited to complete a survey and a Public Goods Game (PGG). Here, we briefly introduce the study setting and methods.

2.1 Study setting

Like many indigenous Oaxacan communities, the village (population 430) is semi-autonomous from the Mexican government, self-governed under a traditional system locally known as *usos y costumbres* ("customs and traditions"). This system involves collective decision-making, communal control of resources, ingroup justice, and community service obligations. As part of this, each male citizen serves a series of unpaid leadership positions called *cargos* (literally, "burdens").

The annual five-day *fiesta* honors the town's patron saint. The celebration involves many instances of collective ritual, including traditional dancing, music-making, processions, and Catholic masses (Figure 1). Although the *fiesta* formally lasts five days, associated preparations and celebrations last for a month and a half (Figure 2, Figure S1). Throughout this time, *cargo*-

holders and other villagers host smaller gatherings that involve dancing and processions, for example to celebrate the donation of bulls to be eaten during the *fiesta*.

Although the *fiesta* is a community endeavor, participation varies. First, while the community is predominantly Catholic, 23% of villagers are Pentecostal Evangelicals (INEGI, 2020). The Evangelicals view the adoration of Catholic saints as idol worship and therefore do not celebrate the *fiesta*. Second, some citizens– including *cargo*-holders, their wives, single mothers, and widows– serve obligatory roles in the *fiesta*. They must provide labor during the *fiesta*, such as coordinating events or working in the communal kitchen. These obligations supersede religious affiliations; Evangelicals must fulfill their duties. Shirking one's duties is punished, typically with a fine.

For a more complete ethnographic description of this community and the *fiesta*, see Supplemental Material S1.

2.2 Participant recruitment and sample size

After an extensive and comprehensive recruitment process involving door-to-door canvasing, 48 participants completed the baseline survey. However, over the course of data collection, 12 participants dropped out (see Supplement S2.1), leaving 36 survey participants (18 female, with a mean age of 44.8). Robustness checks indicate that participant attrition was not biased and therefore unlikely to impact the study's results (S3.7). Of these participants, 34 also completed a post-*fiesta* survey module about *fiesta* participation. A subsample of 27 participants (15 female, mean age of 40) completed both waves of the Public Goods Game. Two of these participants were removed from the PGG analysis because they got fewer than 70% of the test questions correct, suggesting poor understanding of the game. This leaves 25 participants in the PGG sample. Two of these participants did not answer the *fiesta*-wave material security

question, meaning that PGG models controlling for this variable have a sample size of 23 in the *fiesta* wave. All participants gave verbal informed consent before each round of data collection. This research received approval from the Harvard University Area Institutional Review Board (Protocol # IRB21-1632).

2.3 Data collection

Data were collected in two waves, baseline (May-June 2022) and *fiesta* (October 2022) (Figure 2). During the *fiesta* wave, survey data were collected during and soon after the *fiesta*, depending on participant availability. Because the Public Goods Game (PGG) took longer and participants were very busy during the *fiesta,* all *fiesta*-wave PGG data were collected after the *fiesta* (between 2-22 days after the *fiesta*, mean = 8 days).

All data were collected in-person by the first three authors, who were living in the community during the study period. Surveys and PGGs were administered in Spanish to Spanish-speaking or fully Spanish-Zapotec bilingual participants. Apart the PGG decision data (coins in envelopes), all data were collected on tablets.

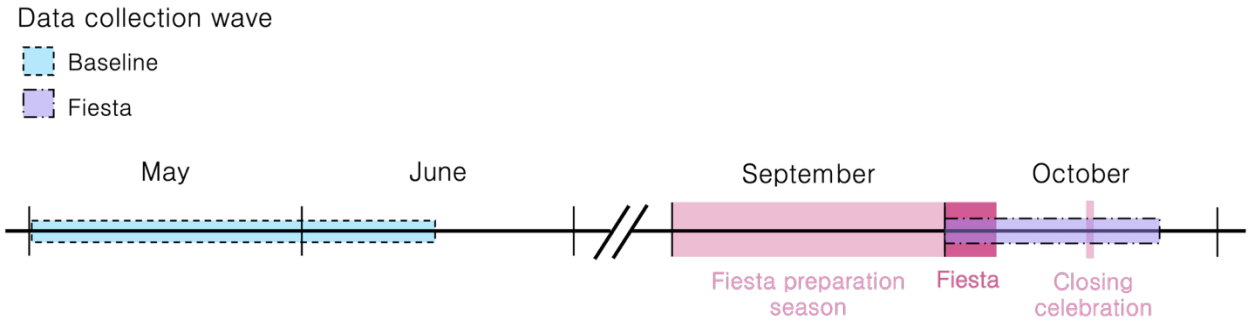

Figure 2. Data collection timeline. Data were collected in two waves, baseline (May-June 2022) and *fiesta* (October 2022). The *fiesta*, which lasted for 5 days in early October, was preceded by about a month of preparatory celebrations and followed by a closing celebration (10 days after the end of the *fiesta*). During the *fiesta* wave, survey data were collected during and soon after the *fiesta*, depending on participant availability. Because the Public Goods Game took longer and participants were very busy during the *fiesta,* all *fiesta*-wave PGG data were collected after the *fiesta* (between 2-22 days after the *fiesta*, mean = 8 days).

2.4 Survey measures

2.4.1 Repeated measures

At both baseline and post-*fiesta*, participants responded to survey questions designed to measure perceptions of group cohesion, prosocial attitudes towards ingroup members, ingroup trust, and household economic stability. For details regarding these measures, see Table 1. See S2.2 for further detail, including information about steps we took to adapt these measures to this field setting, such as the use of visual Likert scales.

2.4.2 Fiesta participation measures

The *fiesta*-wave of data collection included survey questions about donations to the *fiesta*, *fiesta*-associated work obligations, and how the participant felt during the *fiesta* (free-response). In addition, for each of the 5 days of the *fiesta* plus one day before and after, we asked participants whether they participated in collective ritual activities: attending mass, participating in a religious procession, participating in a general procession, playing music in the band, dancing traditional dances (*jarabes*), and performing a dance for an audience (*danzas*). We also asked about participation in other *fiesta* activities, such as doing communal labor or watching dance performances. A Collective Ritual Index was created by summing and then standardizing counts of participation in mass, processions, music-making, and dancing. Because people may have participated in a single activity more than once on a given day, these totals likely underestimate participation.

Table 1. Repeated survey measures

Variable	Category	Description	Source
Willingness to Help (Generalized Altruism)	Ingroup Altruism	"I would help a fellow community-member, even if they could never return the help" (5-point response scale from "strongly disagree" to "strongly agree")	
Likelihood of donating to sick child	Ingroup Altruism	Responding to a vignette, participants reported their likelihood of donating to the family of a sick child for hospital bills (5-point response scale from "very unlikely" to "very likely").	
Likelihood of supporting widow	Ingroup Altruism	Responding to a vignette, participants reported their likelihood of bringing food or other support to a widow after her husband's death (5-point response scale from "very unlikely" to "very likely").	
Identity fusion	Cohesion	Measure of porousness between one's personal and collective identities and visceral feeling of "oneness" with the group. We used the Visual Identity Fusion Scale (Fig. 3).	Adapted from Swann et al. (2009), following Purzycki & Lang (2019) 9/25/26 10:00:00 AM
Shared Fate	Cohesion	Measure of perceived interdependence. Scale includes items like, "When my community feels good, I feel good" and "My community and I rise and fall together" (5-point response scale from "strongly disagree" to "strongly agree"). To create the scale, we averaged across 5 items (Cronbach's $\alpha = 0.79$ [95% CIs 0.67, 0.86]).	Ayers et al. (2023)
Trust	Ingroup Trust	"How much do you trust people in this community?" (5-point response scale from "trust completely" to "distrust completely")	
Material security	Economic stability	"How certain are you that you will be able to buy or produce enough food to eat in the next month?" (5-point response scale from "very uncertain" to "very certain")	Purzycki et al. (2016)

Notes: These measures were collected in both waves of the survey: baseline and post-*fiesta*. For more details about these measures, please refer to the Supplemental Materials (S2.2).

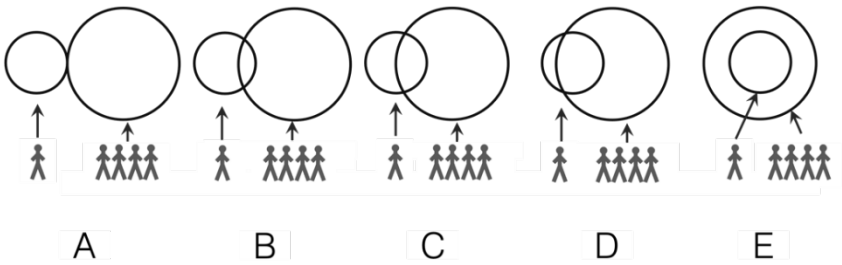

Figure 3. Visual Identity Fusion Scale (adapted from Swann et al., 2009). Participants were asked: "In each of these images, imagine that you are the little circle, and this community is the big circle. Using these pictures, how emotionally close do you feel to this community?"

2.5 Public Goods Game

To measure cooperative behavior, participants played an anonymous, one-shot Public Goods Game (PGG) with three other community members. In this game, each player was asked to divide a cash endowment of $100 MXN (ten $10 MXN coins, about $5 USD total) between two envelopes: a "self" envelope and a "common fund" envelope (Figure S3). Players kept any money that they placed in the "self" envelope, while contributions to the common fund envelope were doubled and then divided evenly between the four players. Payouts from both envelopes were made in a lump sum several weeks after participating. In this classic behavioral economics game, contributions to the common fund are understood as a measure of cooperation. The group can do best if everyone contributes to the common fund; however, an individual can do best if they free ride while the rest of the group cooperates. For more details about the PGG, see Supplemental Materials S2.3.

2.6 Data analysis

We used Linear Mixed Effects Regressions (LMER, R package "lme4", version 1.1–23) to analyze the results. In the basic analysis, for participant *i*:

$$Y_i = \beta_0 + \beta_1 Fiesta_i + \sigma_i + \varepsilon_i$$

where:

- Y one of the following outcome variables: identity fusion, Shared Fate Index, trust, willingness to help, likelihood of donating to the family of a sick child, likelihood of bringing support to a new widow, or PGG allocation
- $Fiesta$ is a categorical variable that equals 0 for baseline and 1 for *fiesta* data collection wave
- σ is the random intercept for participant i
- ε is the error term

Changes in economic stability between waves could impact cooperation in the PGG and willingness to help needy ingroup members. To avoid omitted variable bias, it is therefore important to control for changes in material security in the altruism and PGG analyses. At the same time, if material security is affected by the *fiesta* itself, including this control would introduce post-treatment bias into the estimates. To balance the opposing threats of omitted variable bias and post-treatment bias, we ran all altruism and PGG models with and without the material security control (PGG: Table 2; altruism measures: Tables S1 & S2). Results reported in the Main Text do not control for material security, which the exception of PGG behavior, since this measure is particularly prone to being impacted by changes in material security.

To increase power in this small sample, we stacked variables that measure similar constructs. To capture altruism, we stacked the two altruism vignette outcomes, creating a dataset with 2 observations per participant at each timepoint. We note that Willingness to Help is not correlated with the vignette measures (Likelihood of donating: r = 0.13, p = 0.28; Likelihood of helping widow: r = 0.06, p = 0.64) although they are correlated with each other (r = 0.37, $p <$ 0.001). This suggests that the abstract Willingness to Help and more concrete, locally-relevant vignette questions may capture distinct dimensions of prosociality. Hence, we did not stack

Willingness to Help with the other two measures. Similarly, we did not stack our two measures of cohesion, identity fusion and Shared Fate Index. These measures are not correlated (r = 0.16, p = 0.19), suggesting that they capture distinct dimensions of cohesion. LMER models analyzing the stacked outcome include crossed random intercepts for participant and outcome measure (R package glmmTMB, version 1.1.3).

Finally, we ran several robustness checks. First, because survey questions used 5-point Likert Scales and responses were often clustered towards one end of the scale (see histograms in S3.8), we reanalyzed those outcomes by fitting Cumulative Link Mixed Models (CLMMs) with a logit link (R package "ordinal", version 2022.11-16). Second, because outliers can unduly impact results in such a small sample, we ran Influence Analysis (R package "influence.ME", version 0.9-10) to identify overly influential participants in the main models. We then re-ran those models with the most influential participants removed (details in S3.1.2). Third, since some participants did not complete the PGG until several weeks after the *fiesta*, we analyzed PGG behavior for the sub-sample of participants who played the game within 1 week of the end of the *fiesta*. Prior research suggests that the psychological impacts of collective ritual can last at least a week, and sometimes up to 3 weeks (Páez et al., 2015), although others have found that effects peak during the ritual (Singh et al., 2020). We limited this robustness check to players who participated within 1 week of the *fiesta* (N = 18). This analysis can be found in S3.5.

Below, we report coefficients, 95% Confidence Intervals (CIs), p-values, and degrees of freedom, formatted as p(df).

2.7 Materials & data availability

Study materials, data, and analysis scripts are publicly available at: https://doi.org/10.17605/OSF.IO/AF75N.

3. Results

3.1 Changes in prosocial psychology from baseline to post-*fiesta*

From baseline to after the *fiesta*, results suggest that ingroup altruism declined. Figure 4 shows that the strongest evidence is for a decline in Willingness to Help. After the *fiesta*, there was a 0.44 standard deviation decline in people's willingness to help a community member who couldn't return the help (full sample: β = –0.44, 95%CI [–0.85, –0.02], $p(35) = 0.04$, Table S1, Column 1; Non-Evangelical subsample: β = –0.58, 95% CI [–1.06, –0.10], $p(26) = 0.02$, Table S2, Column 1). There is likewise evidence of declines in the vignette measures of ingroup

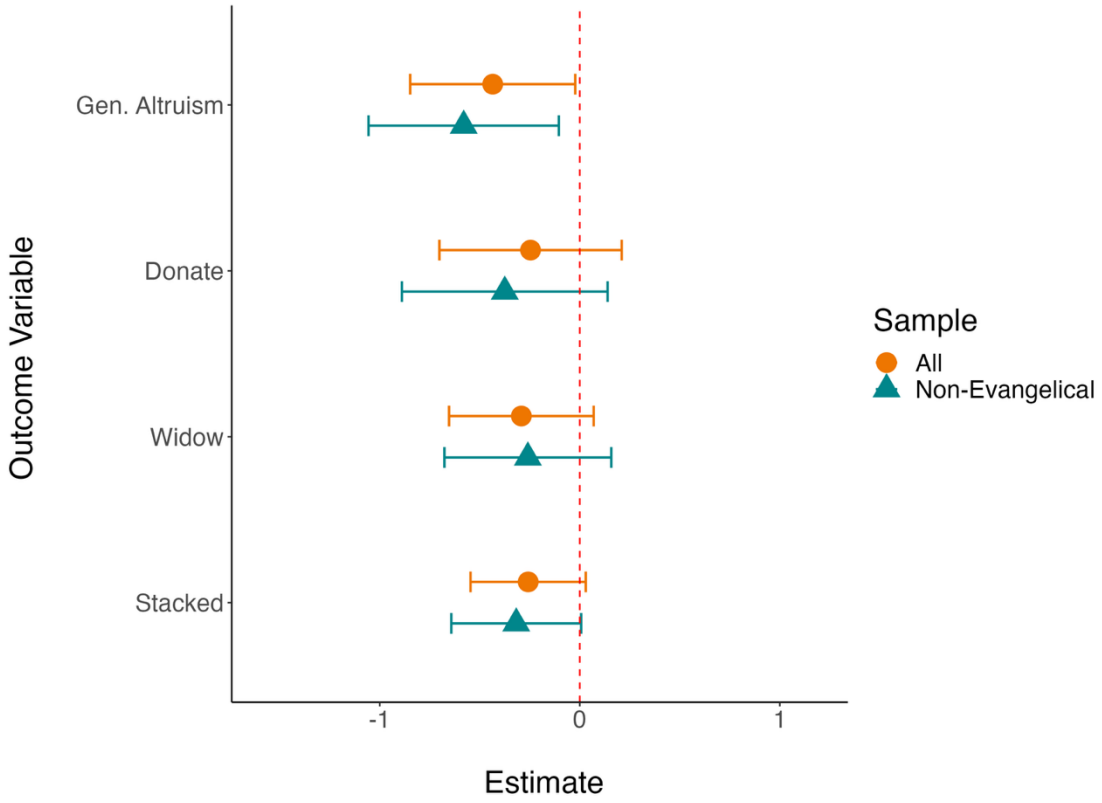

Figure 4. Evidence of declines in ingroup altruism from baseline to post-*fiesta*. This plot shows LMER estimates with 95% confidence intervals for the full sample (36 participants) and the non-Evangelical subsample (27 participants). *Generalized altruism* asked participants to state their level of agreement with the statement, "I would help a fellow community-member, even if they could never return the help." *Donate to a sick child* and *Bring food to widow* represent responses to vignettes wherein participants were asked how likely they would be to help a community member in need. *Stacked* shows the analysis of a stacked dataset, wherein the two vignette-based measures of altruism are collapsed into a single column (4 observations per participant). All continuous variables have been standardized. All models include random intercepts for participant. The stacked models include crossed random intercepts for participant and altruism measure.

altruism, although there is more uncertainty in these estimates. After the *fiesta*, we observed declines in the reported likelihood of bringing food or other support to a new widow (β = –0.29, 95% CI [–0.65, 0.07], $p(35) = 0.11$, Table S1, Column 5) and of donating to the family of a sick child (β = –0.25, 95% CI [–0.70, 0.21], $p(35) = 0.28$, Table S1, Column 3), although the 95% confidence intervals are wide and include 0. When these two vignette measures are stacked (4 observations per participant), there is evidence of a moderate decline in altruism in both the full sample (β = –0.26, 95% CI [–0.55, 0.03], $p = 0.08$; Table S1, Column 7) and non-Evangelical subsample (β = –0.32, 95% CI [–0.64, 0.01], $p = 0.05$; Table S2, Column 7), although the 95% CIs on these coefficients just include 0.

When the most influential participants are removed, evidence of declines in altruism remains. In many cases, influential participants were pulling the coefficients on measures of altruism towards 0. So, removing them tends to result in larger negative coefficients that are more precisely estimated (see S3.1.2, Tables S3 & S4 for details). Analyses using mixed effects ordinal logistic regression (CLMMs with a logit link, more appropriate for Likert scale responses) yielded very similar results to the linear models (details in S3.1.1).

There is no strong evidence of a change in group cohesion from baseline to post-*fiesta*, as measured by identity fusion (β = 0.08, 95% CI [–0.21, 0.36], $p(31.3) = 0.58$) or Shared Fate Index (β = –0.17, 95% CI [–0.54, 0.19], $p(35) = 0.35$). Finally, there was flatly no change in trust from baseline to post-*fiesta* (β = –0.01, 95% CI [–0.38, 0.37], $p(34.4) = 0.97$).

3.2 Changes in cooperation in the Public Goods Game from baseline to post-*fiesta*

From baseline to post-*fiesta*, evidence suggests that contributions to the common fund in the Public Goods Game declined. Controlling for changes in material security, contributions to the common fund were $14 MXN smaller after the *fiesta* (β = –13.89, 95% CI [–28.34, 0.56],

$p(24.9) = 0.06$, Table 2, Column 2), although there is uncertainty in the estimate, with the 95% CIs just including zero. Contributions to the common fund declined from a mean of about \$54MXN at baseline to about \$41MXN after the *fiesta*. Results are similar in the non-Evangelical subsample ($\beta = -15.71$, 95% CI [–32.21, 0.78], $p(20.6) = 0.06$, Table 2, Column 4). In the full sample, Influence Analysis revealed two influential participants, who pull the coefficient in opposite directions. When these participants are removed, the decline in contributions to the common fund is more precisely estimated ($\beta = -14.64$, 95% CI [–26.36, –2.92], $p(21.8) = 0.02$, Table S5, Column 4). Finally, results from the sub-sample of 18 participants who completed the *fiesta*-wave PGG within 1 week of the *fiesta* show similar evidence of declines in contributions to the common fund ($\beta = -16.41$, 95% CI [–33.63, 0.80], $p(16.6) = 0.06$).

Table 2. Evidence of a decline in cooperation in the Public Goods Game

	Contributions to the Common Fund ($MXN)			
	Full sample		Non-Evangelical	
	(1)	(2)	(3)	(4)
Fiesta	–12.00 (–25.48, 1.48)	–13.89+ (–28.34, 0.56)	–13.81 (–29.54, 1.92)	–15.71+ (–32.21, 0.78)
Material security		4.98 (–4.15, 14.11)		5.26 (–3.84, 14.39)
Random Eff.				
Sub.Std. Dev.	12.17	11.92	6.97	6.43
Obs(*N*)	50(25)	48(25)	42(21)	41(21)

Notes: LMER estimates with 95% confidence intervals. *Contributions to the Common Fund* shows how many pesos of their \$100MXN endowment participants contributed to the common fund in the Public Goods Game. *Material security* asked participants how certain they were that they would be able to "buy or produce enough food to eat in the coming month". This variable has been standardized, with higher values indicating greater certainty. The columns labeled "Non-Evangelical" show results from the sub-sample of non-Evangelical participants (Catholics and non-religious). All models include random intercepts for participants. The sample size row shows the number of observations, with the number of participants in parentheses.

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.07$

In contrast, material security increased from baseline to post-*fiesta*: after the *fiesta*, participants reported feeling 0.30 standard deviations more certain that they could buy or produce enough food in the coming month (β = 0.30, 95% CI [–0.01, 0.62], $p(22.4) = 0.06$). In the discussion section, we will discuss how this change likely reflects seasonality.

3.3 Participation in the *fiesta*

On average, Catholics and non-religiously affiliated participants took part in 15.5 *fiesta* activities, including 8.4 collective ritual activities (dancing, processing, music-making, and attending mass) (Figure 5). 54% of non-Evangelicals who had no *fiesta*-related obligation still did communal work on at least one day of the *fiesta*. Evangelicals who had a *fiesta*-related

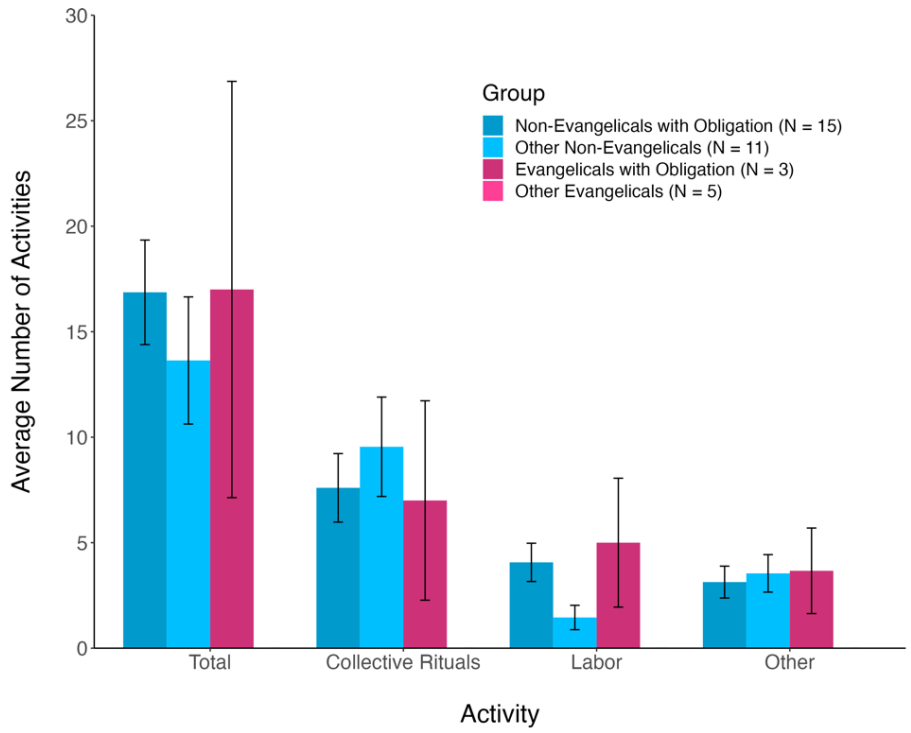

Figure 5. Average number of *fiesta*-related activities participated in during the *fiesta*. Participants are subset by religion and by whether they had an obligation to do *fiesta*-related work (for example, *cargo* duties that required them to serve food in the community kitchen; coordinate dance performances; or ensure safety at the rodeo). Non-Evangelicals include Catholics and those with no religious affiliation. To estimate participation, we asked participants whether they had participated in various activities on each day of the *fiesta* plus one day before and one day after (0 = no, 1 = yes). "Collective ritual" activities include dancing *jarabes*, performing a *danza*, processing, attending mass, and playing music in the band. "Labor" activities include working in the community kitchen and other collective labor associated with the *fiesta*. "Other" activities include amusements such as watching the rodeo, dance performances, basketball tournament, or sparkler show.

obligation ($N = 3$) participated in an average of 17 *fiesta* activities, including 7 collective ritual activities, whereas Evangelicals who had no obligation ($N = 5$) did not take part in any *fiesta*-related activities.

In addition to a $500MXN cash contribution required of all households to help fund the *fiesta*, 50% of Catholic and non-religiously affiliated participants made a voluntary donation, while none of the Evangelical participants donated. Voluntary donations– which included cash, food, costumes for the dances, and decorations– ranged in value from $200MXN (~$10USD, a little less than local average daily wage) to $10,000MXN (~$500USD), with an average value of $2,755MXN (~$135USD). When asked why they donated, the top reasons that participants gave were that: they simply wanted to help; it is local custom to support the *fiesta*; and they had made a vow to a saint to donate something.

When asked how they felt during the *fiesta*, participants gave mixed responses (Table 3). About 44% of participants reported feeling positive emotions, including happiness, pride, and excitement. Meanwhile, about 35% of participants reported feeling negative emotions during the *fiesta*, including stress and fatigue associated with *fiesta*-related responsibilities and annoyance at the poor organization of the *fiesta* by the town leaders. The remaining participants reported neutral emotions (e.g. “normal” and “same as always”).

3.4 Exploratory analyses of moderating effects of *fiesta* participation and *fiesta* emotions

Given the small sample sizes, any analyses with interactions must be considered exploratory. First, to test whether actual participation in collective ritual activities had any moderating effects on psychological outcomes, we included a two-way interaction between data collection wave and Collective Ritual Index in the model. There are no clear patterns of

Table 3. Emotions during the *fiesta* were mixed

Valence	% of Sample	Examples
Negative	35%	• "Stressed due to my obligations [as a *cargo*-holder]." • "Stressed, and tired feet." • "Hectic. Worried about the *fiesta* turning out well." • "Angry that [the leaders] didn't organize the *fiesta* well."
Neutral	21%	• "Calm, passing the day" • "Normal, normal daily life" • "Same as always"
Positive	44%	• "Happy to see my countrymen return and to be convivial with my neighbors." • "Excited that the *fiesta* happened again [after COVID]; it's wonderful that people are convivial, and that people come from other communities." • "Happy, proud to be serving all the people, the townspeople, the visitors." • "Happy, because the *fiesta* was really wonderful."

Notes: Table summarizes responses to the free-response question, "How did you feel during the *fiesta*?" Responses were coded for emotional valence (negative, neutral, or positive).

interaction on measures of ingroup altruism or cohesion in either the full sample or the non-Evangelical subsample (for full results, see Supplemental Materials S3.2).

Second, after discovering that participants experienced a wide range of emotions during the *fiesta* (Table 3), we explored whether emotional valence had any moderating effect on outcomes. Prior research has found that emotional experience can play an important role in the psychological and behavioral impacts of collective ritual (Whitehouse et al., 2017; Whitehouse & Lanman, 2014). Therefore, we included a two-way interaction between emotional valence and data collection wave in the models. Results reveal evidence of a negative interaction between data collection wave and valence of *fiesta* emotions among non-Evangelicals on identity fusion ($\beta = -0.74$, 95% CI [–1.33, –0.15], $p(19.1) = 0.02$) and Shared Fate Index ($\beta = -0.86$, 95% CI [–1.89, 0.17], $p(23) = 0.10$). As illustrated in Figure 6, non-Evangelical participants who reported positive emotions showed signs of increases in identity fusion and Shared Fate Index, whereas those who reported negative emotions may experience declines in these measures of cohesion. In

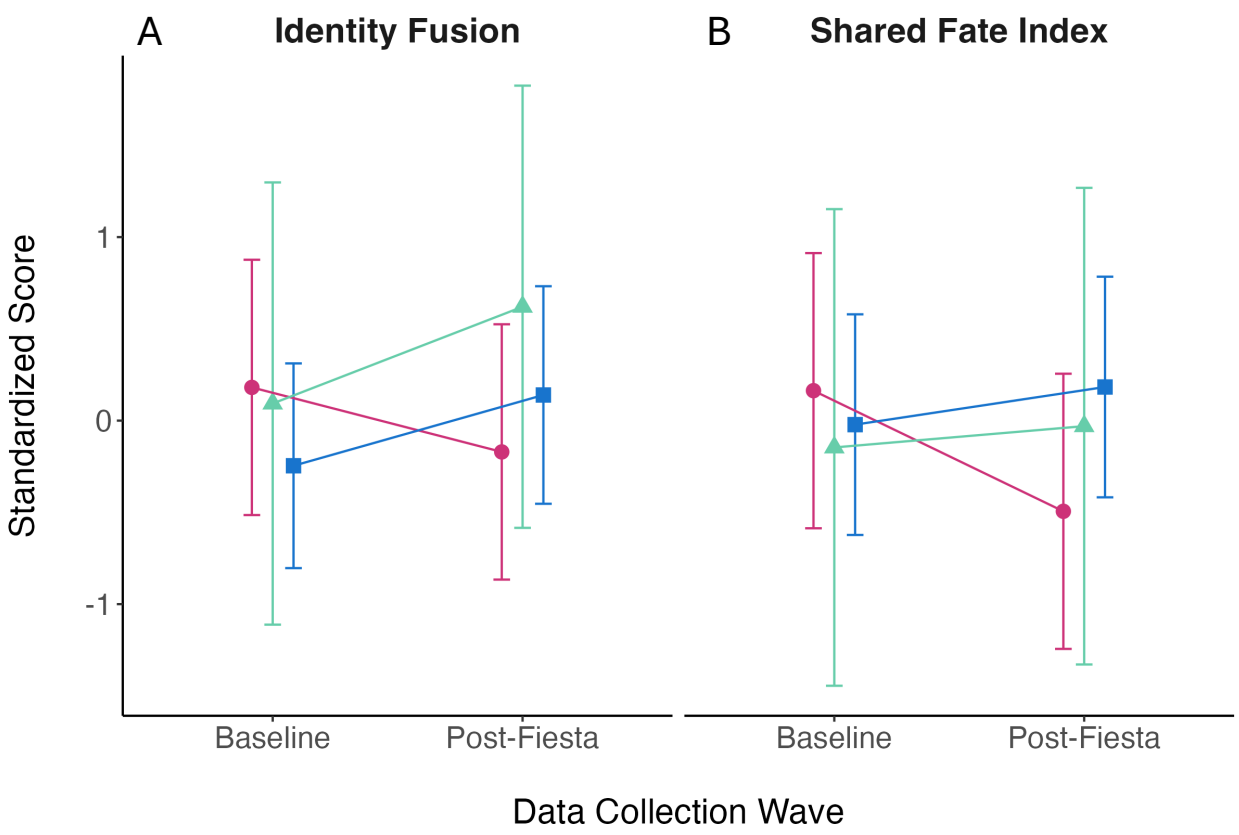

Figure 6. Interactions between *fiesta* emotions and data collection wave on (A) identity fusion and (B) Shared Fate Index. The regression lines show fitted values and 95% confidence intervals produced by LMER models predicting identity fusion and Shared Fate Index from the interaction between data collection wave and valence of emotions felt during the *fiesta*. The models include random intercepts for participants.

the full sample, the coefficient on the interaction term remains negative for identity fusion (β = –0.48, 95% CI [–1.18, 0.21], $p(27.5) = 0.16$) and Shared Fate Index (β = –0.64, 95% CI [–1.50, 0.23], $p(31) = 0.14$), but the coefficients are more imprecisely estimated and the 95% confidence intervals include 0. There is little evidence of any interaction between data collection wave and emotional valence on other outcomes (Supplemental Materials S3.3).

4. Discussion

Our results do not support the hypothesis that cohesion and cooperation increased in the community after the *fiesta*, a large and complex collective ritual. From baseline to post-*fiesta*, there were no changes in cohesion or ingroup trust. Moreover, there is evidence of a decline in

ingroup-oriented altruism and cooperation. These results are not driven by Evangelical participants who disapprove of the *fiesta*; the patterns hold even when Evangelical participants are excluded from the analyses. These findings contrast prior research in both lab and naturalistic settings, which has consistently found evidence that collective rituals enhance group solidarity and prosociality (Jackson et al., 2018; Kirschner & Tomasello, 2010; Páez et al., 2015; Reddish et al., 2013; Wen et al., 2016, 2020; Xygalatas et al., 2011, 2013).

Although we did not find the expected increases in cohesion and cooperation, we found evidence that the *fiesta* demands substantial cooperation within the community. The *fiesta* was a huge cooperative endeavor, involving both obligatory and voluntary cooperation. *Fiesta*-related norms required community members to contribute their time, energy, and money to ensuring that the *fiesta* was a success. In addition to obligatory contributions, we found that half of the non-Evangelical participants made a voluntary donation, and the majority provided voluntary labor. Moreover, ethnographic observations revealed a custom of hospitality during the *fiesta*, wherein villagers prepared extra food at home and welcomed friends and neighbors to eat and socialize. Finally, the lengthy *fiesta* season (Figure 2, Figure S1) featured many smaller celebrations, each of which involved a mutual support institution called *gozona* (Curtin et al., 2026). Villagers supported the host by contributing goods (cash, food items) and labor, which the host will slowly reciprocate over the coming years. Together, this evidence suggests that the *fiesta* does encourage cooperation within the community.

Why, then, do we see an apparent decline in altruism and cooperation after the *fiesta*? One hypothesis is that community members experience cooperation fatigue after such an intense period of cooperation. Villagers may feel that they have done enough for the community during the *fiesta* season and are thus less willing to help or cooperate with ingroup members. In

interpreting low Dictator and Ultimatum Game offers among the Hadza hunter-gatherers, the anthropologist Marlowe (2004) raised a similar possibility. He suggested that food-sharing within Hadza camps led to "weariness", and that low offers reflected an attempt to escape from constant sharing. A similar dynamic could be at play after the *fiesta*, as participants seek reprieve after an intense period of cooperation. A complementary hypothesis is that normative expectations about cooperation change following the *fiesta*. There may be a community-wide understanding that everyone deserves a break, and cooperative norms weaken for a period. Given that community members are *obliged* to provide a large portion of the labor and financing for the *fiesta* (and may be punished for failing to contribute), a third alternative is that intrinsic prosocial motivation is crowded out by external forces after the *fiesta*. Extrinsic motivation– in the form of reward or punishment– can weaken intrinsic motivation (Frey & Jegen, 2001). Finally, it is possible that the religious divisions within the community contributed to the observed effects. Our analyses of the non-Evangelical subsample confirm that the declines in prosociality are not driven by Evangelical participants. However, the presence of Evangelicals who either did not participate or unwillingly participated (e.g. due to obligations as *cargo*-holders) may have impacted the attitudes of non-Evangelical participants.

The main limitation of this study is the small sample size. Although the within-subject design makes the most of the data, the fact remains that analyses with such a small sample warrant caution. Most of the effects reported in this paper are imprecisely estimated, with 95% confidence intervals that include 0. Moreover, the effect sizes are probably inflated (the "winner's curse", Button et al., 2013), meaning that we cannot make strong claims about the magnitude of the declines in altruism or cooperation.

Nonetheless, we believe that these data deserve attention. It is notable that we consistently observe declines in all four measures of ingroup altruism and cooperation, across all model specifications. The sensitivity checks show that the observed declines are not driven by outliers; in fact, when the most influential participants are removed, estimates tend to become more negative and more precisely estimated. Together, these results suggest a real signal, although future studies with larger samples should further probe the psychological and behavioral effects of Oaxacan patron saint festivals.

Another limitation of this correlational study is that changes from baseline to post-*fiesta* may be caused by an unobserved variable. To account for the most obvious alternative– changes in economic stability over the study period– we showed that our results are robust to controlling for material security. This is an important control, since we would expect that willingness to help a community member or contribute in the PGG would decrease if economic stability worsened. There was a small increase in inflation in Mexico during the study period (INEGI, 2023), and prices of some staple goods in the region increased. Interestingly, we found that material security *increased* over the study period. We interpret this as reflecting the local agricultural cycle. Baseline data were collected near the end of the dry season, before crops could be planted, whereas the post-*fiesta* period was very close to the corn harvest (staple crop). However, because our measure of material security did not capture actual "cash on hand", it is still possible that *fiesta*-related expenditures drove the observed declines in prosociality.

Given that several studies have found a positive relationship between collective ritual and group cohesion (Jackson et al., 2018; Páez et al., 2015; Power, 2018), the lack of change in cohesion from baseline to post-*fiesta* is puzzling. It is possible that the null effect is an artifact of low power. However, our exploratory analyses point to a moderating effect of emotional valence

on cohesion: participants who reported feeling positive emotions during the *fiesta* perceived a greater increase in cohesion compared to those who felt negative emotions. While some researchers have proposed that rituals characterized by high levels of *dysphoria* have a potent effect on cohesion (Whitehouse & Lanman, 2014), these results suggest that positive emotions during ritual can better bond groups in some cases.

This study points to new avenues for research into the social and psychological effects of *fiestas*. If they do not enhance cohesion and cooperation, why do costly Oaxacan *fiestas* persist? One possibility is that *fiestas* do foster group solidarity, but the effects emerge over the longer-term– either after a post-*fiesta* dip, or ontogenetically as people grow up participating in many *fiestas.* As a preliminary step, future research should test whether Oaxacan communities with more *fiestas* are more cohesive and cooperative. Second, *fiestas* may shore up networks of social support within the community via mutual support (*gozona*). *Fiestas* provide an important context for ritual exchange and mutual support in Oaxacan communities (Beals, 1970; Ramos Gil, 2017; Stephen, 2005). Social network research should probe whether *fiestas* enhance the cohesion of social support networks (Power, 2018). Finally, *fiestas* may improve or sustain peaceful intergroup relationships in a region where intergroup conflict is common (Cook, 2014; Dennis, 1987; López-Bárcenas, 2004; Yannakakis, 2008). During the *fiesta*, outsiders from nearby communities visit and partake in the festivities. The *fiesta* may function as a credible signal of the community's ability to cooperate on a large-scale (Mehr et al., 2021). Future work should examine the psychological impact of these rituals on visitors from nearby communities.

Broadly, this research adds nuance to the literature on collective ritual. To our knowledge, this is the first study to document declines in altruism and cooperation after a collective ritual. We also found that, despite dancing, music-making, and processing, participants

on average experienced no increase in perceived cohesion. These results suggest that not all collective rituals have psychological effects on cohesion or ingroup-oriented prosociality. Future research should investigate the temporal dynamics of the psychological impacts of collective rituals and consider moderating effects of emotional valence. While the sample is small, we hope that this study will motivate more research into complex, real-world collective rituals and how they shape social psychology in societies around the world.

5. Acknowledgements

We wholeheartedly thank the entire study community for welcoming us, with a special acknowledgement of the participants and the 2022 Authorities for their support. We are grateful to Nathan Nunn, Martin Surbeck, Martin Lang, and three anonymous reviewers for providing valuable feedback on earlier versions of this manuscript. We thank Josh Cetron of Harvard's Institute for Quantitative Social Science for his statistical guidance. This research was supported by the Harvard University Committee on General Scholarships (Sheldon Traveling Fellowship) and the Harvard University Department of Human Evolutionary Biology.

6. Declarations

All authors declare no conflicts of interest.

7. Declaration of generative AI use

During the preparation of the manuscript, Claude Sonnet (v4.6) was used for minor assistance with coding in R. The first author has reviewed and verified the code.

8. References

Ayers, J. D., Sznycer, D., Sullivan, D., Beltrán, D. G., van den Akker, O. R., Muñoz, A. E., Hruschka, D. J., Cronk, L., & Aktipis, A. (2023). Fitness interdependence as indexed by shared fate: Factor structure and validity of a new measure. *Evolutionary Behavioral Sciences*, *17*(3), 259–284.

Beals, R. L. . (1970). Gifting, Reciprocity, Savings, and Credit in Peasant Oaxaca. *Southwestern Journal of Anthropology*, *26*(3), 231–241.

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365–376. https://doi.org/10.1038/nrn3475

Cook, S. (2014). Mexico Santa Cecilia Jalieza: Defending Homeland in Hostile Surroundings. In *Land, Livelihood, and Civility in Southern Mexico* (pp. 190–213). University of Texas Press.

Curtin, C. M., Vasquez Martinez, N., Vásquez Vásquez, Y., & Henrich, J. (2026). *Zapotec cooperative institutions: The psychological and social mechanisms* [Unpublished manuscript].

Curtin, C. M., Vásquez, Y. V., Vásquez Martinez, N., & Henrich, J. (2024). *Traditional political institutions, cooperation, and the provisioning of public goods in Oaxaca, Mexico*.

de la Fuente, J. (1949). *Yalálag: Una Villa Zapoteca Serrana*. Museo Nacional de Antropología.

Dennis, P. A. (1987). *Intervillage Conflict in Oaxaca*. Rutgers University Press.

Durkheim, E. (2001). *The elementary forms of religious life* (C. C. (Trans.), Ed.). Oxford University Press.

Frey, B. S., & Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys*, *15*(5), 589–611. https://doi.org/10.1111/1467-6419.00150

Henrich, J. (2009). The evolution of costly displays, cooperation and religion. Credibility enhancing displays and their implications for cultural evolution. *Evolution and Human Behavior*, *30*(4), 244–260. https://doi.org/10.1016/j.evolhumbehav.2009.03.005

Hobson, N. M., Schroeder, J., Risen, J. L., Xygalatas, D., & Inzlicht, M. (2018). The Psychology of Rituals: An Integrative Review and Process-Based Framework. *Personality and Social Psychology Review*, *22*(3), 260–284. https://doi.org/10.1177/1088868317734944

Instituto Nacional de Estadística Geographía e Informatica. (2020). *Censo de Población y Vivienda*. Instituto Nacional de Estadística, Geografía, e Informática. https://www.inegi.org.mx/programas/ccpv/2020/?ps=microdatos#Datos_abiertos

Irons, W. (2001). Religion as a Hard-to-Fake Sign of Commitment. In R. M. Nesse (Ed.), *Evolution and the Capacity for Commitment*. Russell Sage Foundation.

Jackson, J. C., Jong, J., Bilkey, D., Whitehouse, H., Zollmann, S., McNaughton, C., & Halberstadt, J. (2018). Synchrony and Physiological Arousal Increase Cohesion and Cooperation in Large Naturalistic Groups. *Scientific Reports*, *8*(1), 1–8. https://doi.org/10.1038/s41598-017-18023-4

Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, *31*(5), 354–364. https://doi.org/10.1016/j.evolhumbehav.2010.04.004

Liénard, P., & Boyer, P. (2006). Whence collective rituals? A cultural selection model of ritualized behavior. *American Anthropologist*, *108*(4), 814–827. https://doi.org/10.1525/aa.2006.108.4.814

López-Bárcenas, F. (2004). Territorios indígenas y conflictos agrarios en México. In *El Otro Derecho* (Issues 31–32, pp. 221–246). ILSA.

Marlowe, F. (2004). Dictators and Ultimatums in an Egalitarian Society of Hunter–Gatherers: The Hadza of Tanzania. In J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr, & H. Gintis (Eds.), *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies* (pp. 168–193). Oxford University Press.

Marshall, J. (Director). (2009). *N/um Tchai: The Ceremonial Dance of the !Kung Bushmen* [Video recording]. Documentary Educational Resources. https://video.alexanderstreet.com/watch/n-um-tchai-the-ceremonial-dance-of-the-kung-bushmen

Martínez, V. L. J. (2013). What we need are new customs. In T. A. Eisenstadt, M. S. Danielson, M. J. B. Corres, & C. S. Polo (Eds.), *Latin America's Multicultural Movements: The Struggle Between Communitarianism, Autonomy, and Human Rights*. Oxford University Press. https://doi.org/10.1093/acprof

Mehr, S. A., Krasnow, M. M., Bryant, G. A., & Hagen, E. H. (2021). Origins of music in credible signaling. *Behavioral and Brain Sciences*, *44*. https://doi.org/10.1017/S0140525X20000333

Mogan, R., Fischer, R., & Bulbulia, J. A. (2017). To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect. *Journal of Experimental Social Psychology*, *72*(April), 13–20. https://doi.org/10.1016/j.jesp.2017.03.009

Norenzayan, A., Gervais, W. M., Willard, A. K., Mcnamara, R. A., Slingerland, E., & Henrich, J. (2016). The cultural evolution of prosocial religions. *Behavioral And Brain Sciences*, *39*, 1–65. https://doi.org/10.1017/S0140525X14001356

Norton, M. (2024). *The Ritual Effect*. Simon & Schuster.

Páez, D., Rimé, B., Basabe, N., Wlodarczyk, A., & Zumeta, L. (2015). Psychosocial effects of perceived emotional synchrony in collective gatherings. *Journal of Personality and Social Psychology*, *108*(5), 711–729. https://doi.org/10.1037/pspi0000014

Power, E. A. (2018). Collective ritual and social support networks in rural South India. *Proceedings of the Royal Society B*, *285*, 20180023.

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

Purzycki, B. G., & Lang, M. (2019). Identity fusion, outgroup relations, and sacrifice: A cross-cultural test. *Cognition*, *186*(November 2018), 1–6. https://doi.org/10.1016/j.cognition.2019.01.015

Ramos Gil, I. (2017). Gozona Y Fandango: Fuentes De Legitimidad De La Alianza Matrimonial En Yalálag, Oaxaca. *Península*, *12*(2), 143–168. https://doi.org/10.1016/j.pnsla.2017.06.007

Reddish, P., Fischer, R., & Bulbulia, J. (2013). Let's Dance Together: Synchrony, Shared Intentionality and Cooperation. *PLoS ONE*, *8*(8). https://doi.org/10.1371/journal.pone.0071182

Richerson, P. J., Baldini, R., Bell, A., Demps, K., Frost, K., Hillis, V., Mathew, S., Newton, E. K., Naar, N., Newson, L., Ross, C., Smaldino, P. E., Waring, T. M., & Zefferman, M. (2016). Cultural group selection plays an essential role in explaining human cooperation: A sketch of the evidence. *Behavioral and Brain Sciences*, *39*, e30.

Ruffle, B. J., & Sosis, R. (2007). Does It Pay To Pray? Costly Ritual and Cooperation. *The B . E . Journal of Economic Analysis & Policy*, *7*(1).

Singh, P., Tewari, S., Kesberg, R., Karl, J. A., Bulbulia, J., & Fischer, R. (2020). Time investments in rituals are associated with social bonding, affect and subjective health: A longitudinal study of Diwali in two Indian communities: Diwali Time investment & Social Bonding. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1805). https://doi.org/10.1098/rstb.2019.0430

Sosis, R., & Alcorta, C. (2003). Signaling, Solidarity, and the Sacred: The Evolution of Religious Behavior. *Evolutionary Anthropology*, *12*(6), 264–274. https://doi.org/10.1002/evan.10120

Stephen, L. (2005). Fiesta: The Gendered Dynamics of Ritual Participation. In *Zapotec Women: Gender, Class, and Ethnicity in Globalized Oaxaca* (pp. 250–281). Duke University Press.

Swann, W. B., Gómez, A., Seyle, C. D., 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*, 995–1011.

Tuzin, D. (1976). *The Ilahita Arapesh: Dimensions of Unity.* University of California Press.

Tuzin, D. (2001). *Social Complexity in the Making: A Case Study Among the Arapesh of New Guinea*. Routledge.

Watson-Jones, R. E., & Legare, C. H. (2016). The Social Functions of Group Rituals. *Current Directions in Psychological Science*, *25*(1), 42–46. https://doi.org/10.1177/0963721415618486

Wen, N. J., Herrmann, P. A., & Legare, C. H. (2016). Ritual increases children's affiliation with in-group members. *Evolution and Human Behavior*, *37*(1), 54–60. https://doi.org/10.1016/j.evolhumbehav.2015.08.002

Wen, N. J., Willard, A. K., Caughy, M., & Legare, C. H. (2020). Watch me, watch you: Ritual participation increases in-group displays and out-group monitoring in children: Ritual and children's group behavior. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1805). https://doi.org/10.1098/rstb.2019.0437

Whitehouse, H. (1996). Rites of terror: Emotion, metaphor and memory in Melanesian initiation cults. *Journal of the Royal Anthropological Institute*, *2*, 703–715.

Whitehouse, H., Jong, J., Buhrmester, M. D., Gómez, Á., Bastian, B., Kavanagh, C. M., Newson, M., Matthews, M., Lanman, J. A., McKay, R., & Gavrilets, S. (2017). The evolution of extreme cooperation via shared dysphoric experiences. *Scientific Reports*, *7*(March), 1–10. https://doi.org/10.1038/srep44292

Whitehouse, H., & Lanman, J. A. (2014). The ties that bind us: Ritual, fusion, and identification. *Current Anthropology*, *55*(6), 674–695. https://doi.org/10.1086/678698

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science*, *20*(1), 1–5. https://doi.org/10.1111/j.1467-9280.2008.02253.x

Xygalatas, D. (2023). Culture, Cognition, and Ritual. In M. J. Gelfand, C.-Y. Chiu, & Y.-Y. Hong (Eds.), *Handbook of Advances in Culture and Psychology*. Oxford University Press.

Xygalatas, D., Konvalinka, I., Bulbulia, J., & Roepstorff, A. (2011). Quantifying collective effervescence: Heart-rate dynamics at a fire-walking ritual. *Communicative & Integrative Biology*, *4*(6), 735–738. https://doi.org/10.4161/cib.17609

Xygalatas, D., Mitkidis, P., Fischer, R., Reddish, P., Skewes, J., Geertz, A. W., Roepstorff, A., & Bulbulia, J. (2013). Extreme Rituals Promote Prosociality. *Psychological Science*, *24*(8), 1602–1605. https://doi.org/10.1177/0956797612472910

Yannakakis, Y. (2008). Witnesses, Spatial Practices, and a Land Dispute in Colonial Oaxaca. *The Americas*, *65*(2), 161–192.

Supplemental materials for:

Prosociality declines after a collective ritual: Evidence from a patron saint festival in Oaxaca, Mexico

S1. Ethnographic background

The study community is a Zapotec village in the mountainous Sierra Norte region of Oaxaca, Mexico. It is a small village of 430 residents (Instituto Nacional de Estadística Geographía e Informatica, 2020), although sizeable migrant communities reside in Los Angeles (California, USA), Mexico City, and Oaxaca City.

Like many indigenous Oaxacan communities, the village is semi-autonomous from the Mexican government, self-governed under a traditional system known as *usos y costumbres* ("customs and traditions"). The *usos y costumbres* system involves decision-making in communal assemblies, communal control of resources, community service obligations, and rules about ingroup justice. The *cargo* system is one notable element of *usos y costumbres*. The *cargo* system consists of a hierarchical ladder of civil-religious posts (*cargos*, literally "burdens") that each male citizen must climb over the course of his adult life. *Cargos* include vital positions related to town leadership, judicial function, public works, public safety, and the Catholic church. Other elements of the *usos y costumbres* system in the village include communal labor for public works (*tequio*) and cash taxes (*cooperación*, lit. "cooperation").

S1.1 The Feast of Saint Francis

Each year, the village celebrates several *fiestas* associated with Catholic saints and holidays. The largest annual *fiesta* honors the town's patron saint, Saint Francis of Assisi. The *fiesta* lasts five days, centered around Saint Francis' feast day of October 4. The celebration involves many instances of collective ritual, including traditional dancing, music-making, processions, and Catholic masses. The *fiesta* also features amusements, such as an intercommunity basketball tournament and a rodeo. Throughout the *fiesta*, three free meals per day are served to visitors and community members in the community kitchen.

S1.1.1 Collective ritual activities

Music and dancing are integral elements of Oaxacan *fiestas*. Every town in the Sierra Norte region has at least one woodwind, brass, and percussion band comprising between approximately 15 and 40 musicians. During a *fiesta*, several bands from neighboring towns visit to play. The bands alternate playing throughout the *fiesta*, from early morning well into the night, accompanying everything from processions and dancing to the rodeo.

There are two types of dancing during a *fiesta*: *jarabes* and *danzas* (Main Text Figure 1). *Jarabes* are long, traditional dances. For several hours each evening of the *fiesta*, the bands alternate playing sets of *jarabes* while people of all ages dance in pairs. While the whole group of dancers is not moving in perfect synchrony, they are all dancing the same steps and moving in a synchronous rhythm. *Jarabes* also happened at other times during the *fiesta*, for example during processions. *Danzas* are traditional, rehearsed dances that are performed in costume for an audience. Dancers rehearse extensively in the month leading up to the *fiesta.* The dances are long, repetitive, synchronous, and often performed multiple times over the course of the *fiesta*.

A *fiesta* features many processions. Religious processions (*procesiones*) involve the movement of the statue of Saint Francis. For example, the statue was moved from its normal place in the Catholic Church to the community kitchen for part of the *fiesta*. General processions (*recorridos*) are any other procession through town, always accompanied by at least one band. Examples include a many-hour procession with dancing that formally opens the *fiesta* (Main Text Figure 1A), the receptions and farewells of the visiting bands, and processions to fetch items that have been donated for the *fiesta* by a family in the community. These processions ranged in size from a couple dozen people to several hundred people, not including the bands.

S1.1.2 The fiesta *season*

Although the *fiesta* itself is five days long, associated preparations and celebrations last for about a month and a half (Figure S1). Formal preparations for the *fiesta* begin in early September with a social gathering (*convivio*) hosted by one the top *cargo*-holders. Throughout September, weekly *convivios* celebrate the donation of bulls to be slaughtered and eaten during the *fiesta*; in 2022, three families donated bulls. Two days before the *fiesta* begins, another top *cargo*-holder hosts a *convivio*. Finally, about eight days after the conclusion of the *fiesta*, a third top *cargo*-holder hosts the Octave– a celebration eight days after a feast day in the Catholic tradition. This formally marks the end of the *fiesta* season. Each *convivio* involves a procession through town with the band, a meal for the community, and some *jarabes*. To help defray the

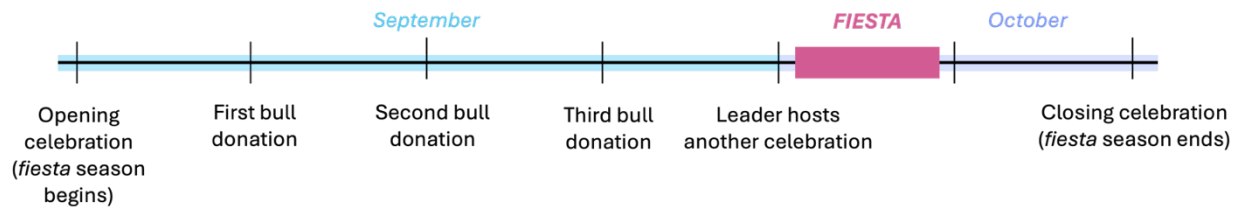

Figure S1. The *fiesta* season. Celebrations and preparations last for approximately 1.5 months, beginning in early September and concluding about 1 week after the *fiesta.* Throughout this time, *cargo*-holders and other villagers host smaller celebrations for the community that include dancing, eating, and processing.

costs of hosting a *convivio*, community members contribute sugar, chocolate, corn, and other food items; labor, such as making tortillas and tamales; and cash. These contributions are made within the context of a mutual aid institution called *gozona*– support that will be reciprocated later.

S1.1.3 Participation in the fiesta

Although the *fiesta* is a community endeavor, participation varies. First, while the community is predominantly Catholic, there is a small Evangelical Pentecostal congregation and a small number of non-religiously affiliated individuals. While non-religious people participate in the *fiesta*, the Evangelicals view the adoration of Catholic saints as idol worship and therefore do not celebrate the town's *fiestas*. Second, some citizens serve obligatory roles in the *fiesta*. For example, some *cargos* involve specific *fiesta*-related duties, such as hosting *convivios*, coordinating events, and managing the communal kitchen. The wives of top *cargo*-holders must work in the communal kitchen throughout the entire *fiesta*, while single mothers and widows must contribute one day of labor. These obligations supersede religious affiliations; Evangelicals must fulfill their duties. For example, in 2022, the most important religious *cargo* position (*mayordomo*) was held by an Evangelical man.

S2. Supplementary methods

S2.1 Participant recruitment, sample size, and data collection timeline

To recruit participants, our team first knocked on the door of every house and invited an adult household member to participate. When this method yielded low participation rates, additional participants were recruited via snowball and convenience sampling. Data were collected in Spanish with fully Spanish-fluent participants. All participants gave informed consent. This research was approved by the Harvard-Area IRB.

Forty-eight participants completed the baseline survey. However, 12 participants dropped out due to no longer wishing to participate (N = 4), being too busy to participate (N = 3), absence from the community (N = 3), or the data collectors being unable to locate the participant (N = 2) (see Supplementary Materials S3.6 for robustness checks regarding participant attrition). Thus, thirty-six participants (18 female, with a mean age of 44.8) completed both waves of the survey. Of those, 34 also completed a post-*fiesta* survey module about *fiesta* participation. A subsample of 27 participants (15 female, mean age of 40) also completed both waves of the Public Goods Game.

Data were collected in two waves, baseline (May-June 2022) and *fiesta* (October 2022). During the *fiesta* wave, survey data were collected during and soon after the *fiesta*, depending on participant availability. Because the Public Goods Game (PGG) took longer and participants were very busy during the *fiesta,* all *fiesta*-wave PGG data were collected after the *fiesta* (between 2-22 days after the *fiesta*, mean = 8 days).

Since some participants did not complete the PGG until several weeks after the *fiesta*, we conducted a robustness check wherein we analyzed PGG behavior for the sub-sample of participants who played the game within 1 week of the end of the *fiesta*. Prior research suggests that the psychological impacts of collective ritual can last at least a week, and sometimes up to 3 weeks (Páez et al., 2015), although others have found that effects peak during the ritual (Singh et al., 2020). We limited this robustness check to players who participated within 1 week of the *fiesta*. This analysis can be found in S3.5.

S2.2 Survey measures

The baseline survey included demographic questions, while both waves included a material security question ("How certain are you that you will be able to buy or produce enough food to eat in the next month?").

Dependent variables

To assess changes in cohesion and prosocial attitudes, we measured several dependent variables at both baseline and post-*fiesta*. We drew on established, validated measures from social psychology where possible. However, as others have documented elsewhere (Hruschka et al., 2018), some changes to these measures were necessary for working in this field setting. For example, to help participants respond to Likert-scale items, we presented visual Likert scales (Fig. S2). As the data collector read the Likert items aloud, they pointed to each option on the visual scale. Participants could respond verbally and/or by pointing to their selection on the scale. Translation and adaptation of questions and scale items to this field setting was a collaborative effort accomplished by the first three authors of this paper, two of whom are young Zapotec scholars from towns nearby the study community with deep insights into the local context and culture. Specific adjustments to key measures are described below.

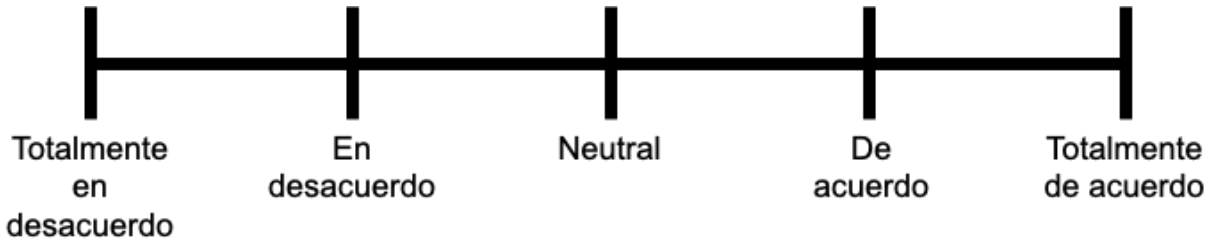

Figure S2. A visual Likert scale. When responding to Likert-type question, participants were presented with a visual scale like this one, which represents a 5-point scale from "completely disagree" to "completely agree". The data collector pointed to each option as they read the items aloud in order. Participants could respond verbally and/or by pointing to their selection on the scale.

Cohesion. We measured identity fusion– a widely-used measure of porousness between one's personal and collective identities and visceral feeling of "oneness" with the group (Gómez et al., 2011; Vázquez et al., 2017). Following (Purzycki & Lang, 2019), we adapted Swann et al.'s (2009) **Visual Identity Fusion Scale** (Main Text Figure 2). We made several adjustments to the visual scale and verbal prompt. First, to enhance the visual clarity of the scale, we included stick figures rather than words to represent "self" (one stick figure) and "community" (group of four stick figures). Second, we included more concrete instructions in the verbal prompt. The original prompt (Swann et al., 2009) is vague and abstract ("Please circle the letter below the picture that best represents your relationship with this group"). Our version of the question gave concrete context for participants: "In each of these images, imagine that you are the little circle and this community is the big circle. Using these pictures, how emotionally close do you feel to this community?"

We also measured perceptions of fitness interdependence, using the **Shared Fate Scale** (Ayers et al., 2023). This 6-item scale asks participants to state how much they agree or disagree with statements like, "When my community feels good, I feel good" and "My community and I rise and fall together." During data collection, we found that many participants struggled with the reverse-coded item ("I don't care whether my community thrives or not"), selecting "agree" or "strongly agree" on the Likert scale, while also making comments like, "No, I *do* care". Therefore, we removed this item from our analysis. The remaining 5 items were highly internally consistent (Cronbach's $\alpha = 0.79$ [95% CIs 0.67, 0.86]), so we averaged them to create a Shared Fate Index. While this construct does not measure cohesion, per se, positive interdependence and cohesion are likely linked: researchers have suggested that identity fusion may be a proximate mechanism that helps humans infer their interdependence with others (Aktipis et al., 2018).

Altruism. As a general measure of Willingness to Help an ingroup member, participants were asked to report their level of agreement with the statement: "I would help a fellow community-member, even if they could never return the help" (5-point response scale from "strongly disagree" to "strongly agree"). Participants also responded to two locally-relevant vignettes featuring situations involving altruism towards community members. Participants were asked how likely it was that they would (1) donate to the family of a sick child for hospital bills and (2) bring food to a widow after her husband's death (5-point response scale from "very unlikely" to "very likely").

Trust. To assess trust in ingroup members, we asked participants: "How much do you trust people in this community?" (5-point response scale from "trust completely" to "distrust completely").

S2.3 Public Goods Game

The Public Goods Game was administered individually in a private setting, away from other community members. Participants were first told how the game works. After a basic

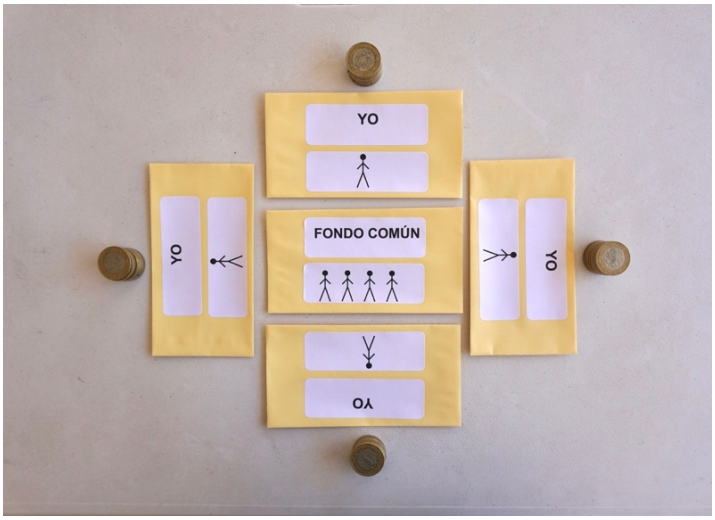

Figure S3. The Public Goods Game. This image shows the basic setup of the game, where each of four anonymous players privately decided how to divide a $100MXN cash endowment between a "self" envelope (labeled "*Yo*", meaning "me") and a "common fund" envelope (labeled "*Fondo Común*"). Contributions to the Common Fund were doubled and then divided evenly between the four players.

overview of the rules, the data collector ran through 5 examples, demonstrating with coins and envelopes (Figure S3) how the game could be played. The participant then responded to a series of test questions to ensure their comprehension. The data collector left the participant alone to make their decision and allocate ten $10MXN coins between the "Self" and "Common Fund" envelopes. To ensure the participant's privacy (so that the data collector could not feel the weight of the coin distribution in the two envelopes), the participant placed both sealed envelopes in a larger manila envelope before the data collector returned. After playing, participants were randomly assigned into anonymous groups of 4 for payout calculation. Payouts were made several weeks later, in a lump sum with winnings from another behavioral economics game for a different study. The second game was an anonymous two-player designed to measure envious behavior. We combined winnings from these two games in order to ensure that participants could not infer the behavior of the other player in the second game.

As part of another study, participants responded to a series of vignettes before playing the PGG. This was part of a priming study run during the baseline wave of data collection. The vignettes consisted of short stories about violations of local social norms, such as beating a dog or disrespecting an elder. Participants were asked to make judgments about the behavior of the protagonist (for example, how bad the behavior was; how bad other people would think it was; and whether he/she should be punished).

In the baseline wave, participants played the PGG twice: once after hearing a set of control vignettes, and once after hearing a set of priming vignettes. The order of presentation of control versus priming vignettes was randomized. We found no effect of this priming manipulation (see Curtin et al., 2026 for full results). To maximize comparability of PGG behavior between waves, participants responded to the control vignettes before playing the PGG

in the post-*fiesta* wave. Thus, in the current study, we compare PGG behavior after exposure to the control vignettes at baseline and post–*fiesta*.

S3. Supplementary results

Complete R code and an R notebook HTML file with all analyses can be found at: https://doi.org/10.17605/OSF.IO/AF75N.

S3.1 Robustness checks

S3.1.1 Cumulative Link Mixed Models

Because survey response options were Likert-scales, we re-analyzed the data using Cumulative Link Mixed Models (CLMMs) with a logit link function. Results revealed the same pattern of declines in altruism as the LMER analyses. When the three measures of altruism are stacked (6 observations per participant), the odds of selecting a more prosocial scale item declined by 70% after the *fiesta* in the full sample (OR = 0.30 95%CIs[0.14, 0.63], p = 0.002; N = 36) and by 68% in the non-Evangelical subsample (OR = 0.32 95%CIs[0.15, 0.71], p = 0.005; N = 27). For example, in the full sample, the probability that participants selected the most prosocial option ("strongly agree" or "very likely") declined from 64% at baseline to 35% after the *fiesta* (this decline was largely absorbed by a proportional increase in the probability of selecting the second most prosocial option, "agree" or "likely"). Analyses also revealed a substantial decline in the odds of selecting a more prosocial scale item after the *fiesta* when only the two vignette measures are stacked in the full sample (OR = 0.43[0.17, 1.12] p = 0.09; N = 36) and non-Evangelical subsample (OR = 0.53[0.20, 1.41] p = 0.20; N = 27), although the 95% CIs are wide and include 1. In the full sample, for example, the probability that participants

stated that they were "very likely" to donate to the family of a sick child or bring support to a widow decreased from 71% at baseline to 51% after the *fiesta*.

When each outcome is analyzed separately, the pattern is similar. After the *fiesta*, the odds of selecting a more prosocial scale item in response to the Willingness to Help question declined by 75% (OR = 0.25 95%CIs[0.08, 0.72], p = 0.01; *N* = 36) in the full sample and 83% in the non-Evangelical subsample (OR = 0.17 95%CIs[0.05, 0.59], p = 0.005; *N* = 27). For example, in the full sample, the probability that participants said they "strongly agreed" that they would help a fellow community member who could never return the help declined from 54.1% at baseline to 22.6% after the *fiesta*. The odds of selecting a higher likelihood of donating money for a sick child's hospital bills (OR = 0.53[0.16, 1.8], p = 0.30) and bringing support to a new widow (OR = 0.29[0.04, 1.99], p = 0.21) likewise declined, although the 95% confidence intervals on these odds ratios are wide and include 1.

S3.1.2 Sensitivity analyses

As an additional robustness check, we ran Influence Analysis (R package "influence.ME", version 0.9-10) to identify overly influential participants who might be driving results. We then re-ran the models with the most influential participants removed. Given our already small sample, we had to be conservative when removing participants; that is, we could not always remove all participants who passed the conventional DFBETA threshold for being considered "influential" ($\frac{2}{\sqrt{n}}$, e.g., 0.33 for a sample of 36). Instead, we applied a consistent rule, removing participants whose DFBETA value for the treatment coefficient was greater than $\frac{2}{\sqrt{n}}$ (the standard threshold), starting with those with the highest DFBETA, up to a maximum of 10% of the sample (e.g., no more than 3 participants removed from a sample of 36). Note that this

approach sometimes left participants with a DFBETA above the standard DFBETA threshold in the model.

In the analysis stacking the two vignette measures of altruism, Influence Analysis identified two influential participants, one with a positive DFBETA and one with a negative DFBETA on the negative *fiesta* coefficient. This indicates that they were pulling the coefficient towards zero– that is, suppressing the negative treatment effect of the *fiesta* on these measures of altruism. When the these participants are removed, the decline in vignette-based altruism measures from baseline to post-*fiesta* is of a similar magnitude as the estimate based on the full sample, but more precisely estimated with 95% CIs just exclude 0 (β = –0.25, 95% CI [–0.48, –0.03], $p = 0.03$; $N = 34$; Table S3, Column 8).

When looking at each measure of altruism separately, Influence Analysis identified three participants who were unduly impacting estimation of the effect of the *fiesta* on Willingness to Help. Two of these participants showed positive DFBETAs on the negative *fiesta* coefficient, indicating that they were pulling the coefficient towards zero. When the three influential participants are removed from the model, the decline in Willingness to Help is larger and more precisely estimated than before (β = –0.56, 95% CI [–0.87, –0.25], $p < 0.001$; $N = 33$; Table S3, Column 2).

In the model analyzing reported changes in likelihood of donating to a sick child, Influence Analysis identified two influential participants. One showed a positive DFBETA on the negative *fiesta* coefficient and one showed a negative value. When these participants are removed, results show still show evidence of a decline in likelihood of donating to a sick child after the *fiesta*, of similar magnitude as before but slightly more precisely estimated (β = –0.22, 95% CI [–0.52, 0.07], $p = 0.14$, $N = 34$; Table S3, Column 4). Similarly, we identified 4

influential participants in the model estimating changes in the likelihood of bringing support to a widow, all of which showed a positive DFBETA on the negative *fiesta* coefficient. When three are removed from the model, we found a larger and more precisely estimated decline than before (β = –0.48, 95% CI [–0.80, –0.16], $p = 0.005$, $N = 33$; Table S3, Column 6).

Influence Analysis of altruism models including the material security control showed very similar patterns: the coefficients on *fiesta* become more negative and more precisely estimated when the most influential participants are removed. See Table S4 for details.

Influence Analysis identified several influential participants in models analyzing changes in cohesion and trust. When the three most influential participants (out of four) are removed from the model analyzing identity fusion, results suggest that there may be an increase in fusion from baseline to post-*fiesta* (β = 0.17, 95% CI [–0.04, 0.37], p = 0.11, $N = 33$). However, when the one influential participant is removed from the model analyzing changes in Shared Fate Index, the coefficient becomes larger and more negative, with 95% CIs that exclude 0 (β = –0.30, 95% CI [–0.56, –0.03], $p(34) = 0.03$). This suggests that, if anything, participants may feel less interdependent with their community after the *fiesta*. Removing influential participants from the model analyzing trust has no appreciable impact on results: we still observe no change in trust over the study period (β = 0.07, 95% CI [–0.21, 0.35], $p = 0.63$, $N = 33$).

S3.2 Exploratory analyses: Testing for moderating effects of Collective Ritual Index

In the full sample, there are no clear patterns of interaction between data collection wave and Collective Ritual Index on the psychological measures of ingroup altruism or cohesion: willingness to help (β = 0.28, 95% CI [–0.15, 0.72], $p = 0.19$), likelihood of bringing food to a new widow (β = –0.16, 95% CI [–0.55, 0.23], $p = 0.41$), Shared Fate Index (β = 0.16, 95% CI [–0.23, 0.56], $p = 0.40$), identity fusion (β = 0.20, 95% CI [–0.18, 0.59], $p = 0.29$), or trust (β =

0.16, 95% CI [–0.25, 0.56], $p = 0.43$). There is no evidence of an interaction between data collection wave and Collective Ritual Index on contributions to the Common Fund in the Public Goods Game (β = 1.36, 95% CI [–17.5, 20.3], $p = 0.88$). We did, however, find evidence of a negative interaction on likelihood of donating to a sick child (β = –0.42, 95% CI [–0.89, 0.05], $p = 0.08$), which would suggest that people who participated in more collective ritual activities saw a *greater* decline in willingness to donate from baseline to post-*fiesta*. Similarly, when the two vignette measures of altruism are stacked, results reveal a negative interaction between data collection wave and Collective Ritual Index (β = –0.30, 95% CI [–0.60, 0.00], $p = 0.05$).

In the non-Evangelical subsample, interacting data collection wave and Collective Ritual Index on outcome variables yields mixed results. There is a positive interaction on willingness to help (β = 0.50, 95% CI [0.03, 0.97], $p = 0.04$). The coefficient on this interaction term represents the change in slope between the data collection wave and willingness to help as participation in collective ritual activities increases. Thus, as participation in collective ritual activities increased, so did the increase in willingness to help from baseline to post-*fiesta*. On the other hand, there is evidence of a negative interaction on likelihood of donating to the family of a sick child (β = –0.44, 95% CI [–0.95, 0.08], $p = 0.09$), although the 95% confidence intervals include 0. This negative interaction suggests that, as participation in collective ritual activities increased, there was less of an increase in likelihood of donating from baseline to post-*fiesta*. There is no evidence of an interaction between data collection wave and Collective Ritual Index on any other outcomes in the non-Evangelical sub-sample: likelihood of bringing food to widow (β = –0.13, 95% CI [–0.58, 0.31], $p = 0.55$), fusion (β = 0.23, 95% CI [–0.19, 0.64], $p = 0.27$, Shared Fate Index (β = 0.17, 95% CI [–0.33, 0.66], $p = 0.50$), or trust (β = 0.16, 95% CI [–0.30, 0.61], $p = 0.48$).

S3.3 Exploratory analyses: Testing for moderating effects of *fiesta* emotions

There is no evidence of an interaction between data collection wave and emotional valence during the *fiesta* on willingness to help (β = –0.64, 95% CI [–1.62, 0.34], $p = 0.19$), likelihood of donating to a sick child (β = 0.63, 95% CI [–0.46, 1.72], $p = 0.24$), bringing food to a new widow (β = –0.06, 95% CI [–0.95, 0.83], $p = 0.89$), or trust (β = –0.16, 95% CI [–1.04, 0.71], $p = 0.70$) in the full sample. When survey measures of altruism are stacked (4 observations per participant), there is no interaction between emotional valence and data collection wave on altruism in the full sample (β = 0.33, 95% CI [–0.35, 1.01], $p = 0.34$) or non-Evangelical sub-sample (β = –0.01, 95% CI [–0.74, 0.72], $p = 0.98$).

S3.4 Exploratory analyses: Testing for moderating effects of *cargo* status

Finally, knowing that *cargo*-holders and their wives provide a large amount of labor and economic input to the community throughout the year and during the *fiesta*, we examined whether *cargo* status moderated any of the outcomes. In the full sample, there is no evidence of an interaction between data collection wave and status as a *cargo*-holder or wife of a *cargo*-holder willingness to help (β = –0.22, 95% CI [–1.11, 0.66], $p = 0.61$), likelihood of donating to a sick child (β = 0.09, 95% CI [–0.89, 1.10], $p = 0.85$), bringing food to a new widow (β = –0.59, 95% CI [–1.35, 0.16], $p = 0.12$), trust (β = 0.23, 95% CI [–0.57, 1.03], $p = 0.56$), identity fusion (β = –0.09, 95% CI [–0.69, 0.51], $p = 0.77$), or Shared Fate (β = 0.16, 95% CI [–0.62, 0.93], $p = 0.68$).

However, there was some evidence of a positive interaction between *cargo* status and data collection wave on PGG contributions (β = 22, 95% CI [–7.21, 51.11], $p = 0.13$), although the 95%CIs are wide and include 0. As illustrated in Figure S5, *cargo*-holders and their wives contributed less than non-*cargo*-holders at baseline, but their contributions of about $37MXN

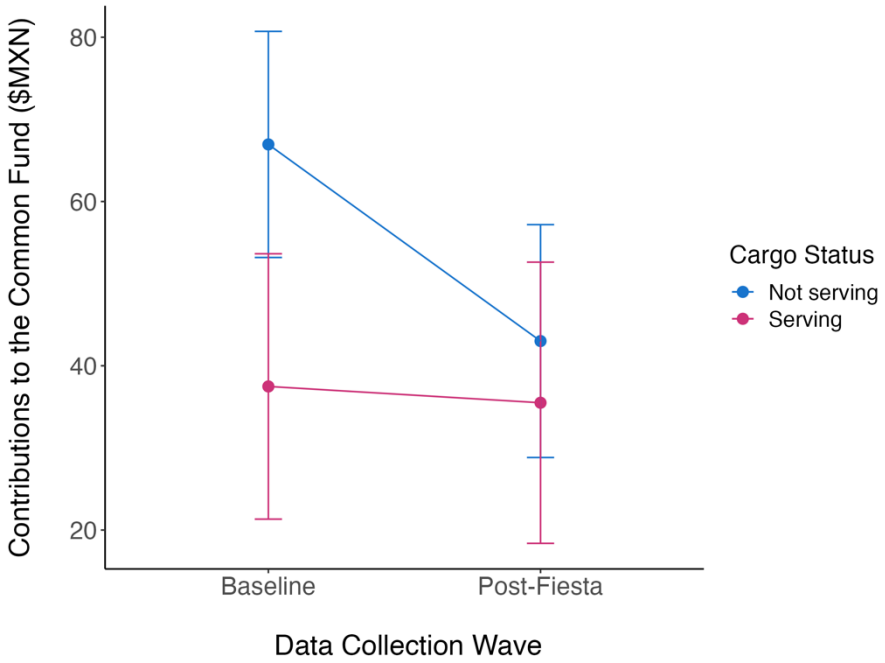

Figure S5. Interactions between *cargo* status and data collection wave on Public Goods Game contributions. The regression lines show fitted values and 95% confidence intervals produced by a LMER model predicting contributions to the Common Fund from the interaction between data collection wave and status as a *cargo*-holder or the wife of a *cargo*-holder. The model includes random intercepts for participants and a control for changes in material security.

remained steady after the *fiesta*. In contrast, non-*cargo*-holders showed a large decline in contributions from baseline to post-*fiesta*, from an average of $66MXN to $44MXN.

S3.5 Public Goods Game behavior in participants who played within 1 week of the *fiesta*

Because some participants did not complete the PGG until several weeks after the *fiesta*, we ran a robustness check, analyzing PGG behavior for the sub-sample of 18 participants who played the game within 1 week of the end of the *fiesta*. Despite the small sample, the results are similar to those of the full sample. Without controlling for material security, participants contributed $16MXN fewer after the *fiesta* compared to baseline ($\beta = -16.11$, 95% CI [–32.48, –0.26], $p = 0.05$, $N = 18$). Once a control for material security is added, the point estimate holds steady, although it is slightly less precisely estimated ($\beta = -16.41$, 95% CI [–33.63, 0.80], $p = 0.06$, $N = 18$).

S3.6 Correlates of Public Goods Game behavior

Apart from our main hypotheses about changes in prosocial attitudes, cohesion, and cooperation from baseline to post-*fiesta*, we also predicted that Public Goods Game allocations would be positively correlated with measures of prosociality and cohesion. To test this hypothesis, we ran LMER with random intercepts for participant, taking advantage of data collected in both waves. We found no evidence of associations between survey measures of prosociality or cohesion and allocations to the Common Fund. In the following results, coefficients represent the change in allocation to the Common Fund in Mexican pesos ($MXN) for a 1 standard deviation change in the predictor: willingness to help (β = –1.63, 95% CI [–12.22, 8.35], p = 0.74), likelihood of donating to the family of a sick child (β = 2.95, 95% CI [–8.05, 13.97], p = 0.60), likelihood of bringing support to a new widow (β = –1.88, 95% CI [–

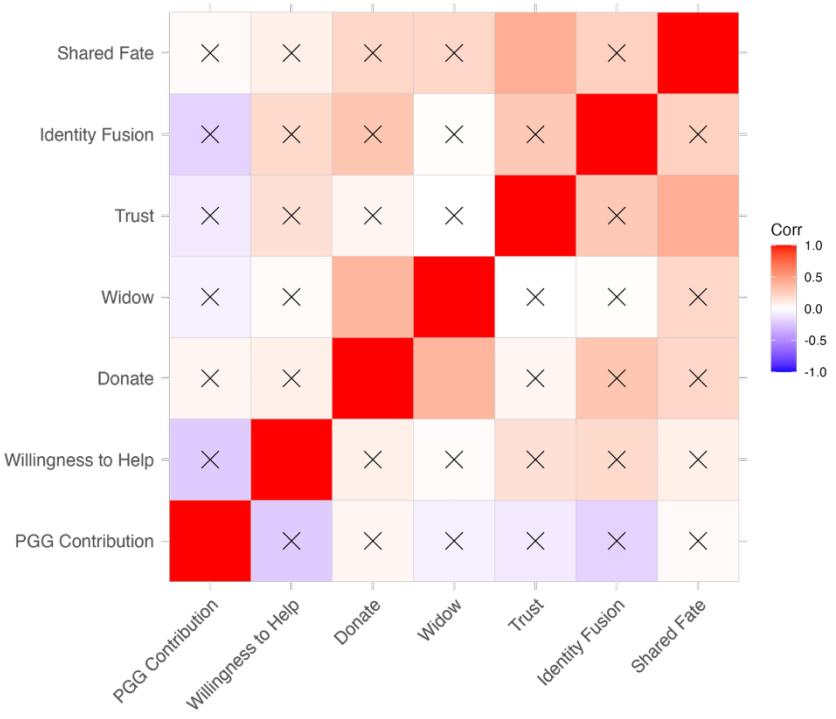

Figure S6. Pairwise Pearson correlations between outcome variables. Correlations with a p-value above 0.05 are marked with an "X".

9.27, 5.52], $p = 0.62$), trust (β = –1.62, 95% CI [–9.56, 6.33], $p = 0.69$), identity fusion (β = –2.43, 95% CI [–10.86, 5.97], $p = 0.57$), or shared fate (β = 0.31, 95% CI [–7.44, 8.11], $p = 0.94$). Pairwise correlations are visualized in Figure S6.

S3.7 Participant attrition robustness checks

To investigate whether participant attrition is likely to have introduced bias into our results, we examined whether the 12 participants who dropped out differed from the rest of the sample in terms of demographics, baseline survey responses, or baseline PGG behavior. Our results offer no evidence that participant attrition introduced bias.

To test whether these samples differ in demographic characteristics, we ran a logistic regression predicting dropout status (binary variable, 1 = dropped out) on participant sex, age, education, religion, material security, *cargo*-status, and the identity of the data collector. None of these variables are significantly associated with dropping out, although results suggest that participants who dropped out may be less educated than the rest of the sample. A one standard increase in level of education is associated with a 58% decrease in likelihood of dropping out of the study (OR = 0.42, 95% CI [0.13, 1.06], p = 0.10), although the confidence intervals are wide and include 1.

To test whether participants who dropped out of the study differed from the rest of the sample in terms of their survey responses or PGG behavior, we regressed each standardized outcome of interest on status as a dropout (binary variable, 1 = dropped out) and the following demographic controls: sex, age, education, religion, material security, *cargo*-status, and the identity of the data collector. We found no evidence that dropping out was associated with outcome measures: willingness to help (β = –0.05, 95% CI [–0.68, 0.57], $p = 0.86$), likelihood of donating to the family of a sick child (β = 0.08, 95% CI [–0.65, 0.81], $p = 0.82$), likelihood of

bringing support to a new widow (β = –0.14, 95% CI [–0.81, 0.54], $p = 0.69$), trust (β = 0.18, 95% CI [–0.55, 0.90], $p = 0.62$), identity fusion (β = 0.49, 95% CI [–0.27, 1.25], $p = 0.20$), shared fate (β = –0.16, 95% CI [–059, 0.90], $p = 0.67$), or contributions to the Common Fund in the Public Goods Game (β = 0.21, 95% CI [–0.38, 0.80], $p = 0.48$).

S3.8 Response variable distribution plots

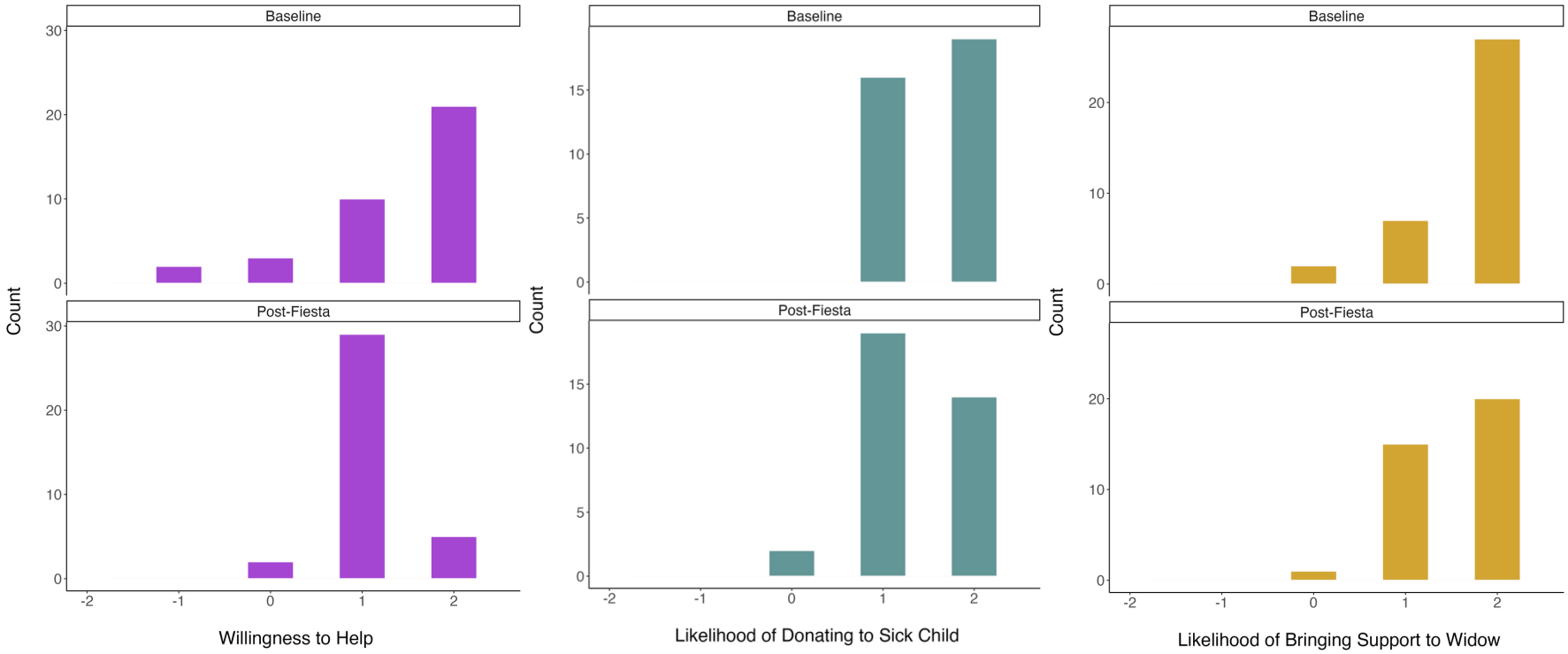

Figure S8. Histograms of survey responses to altruism questions before and after the *fiesta*.

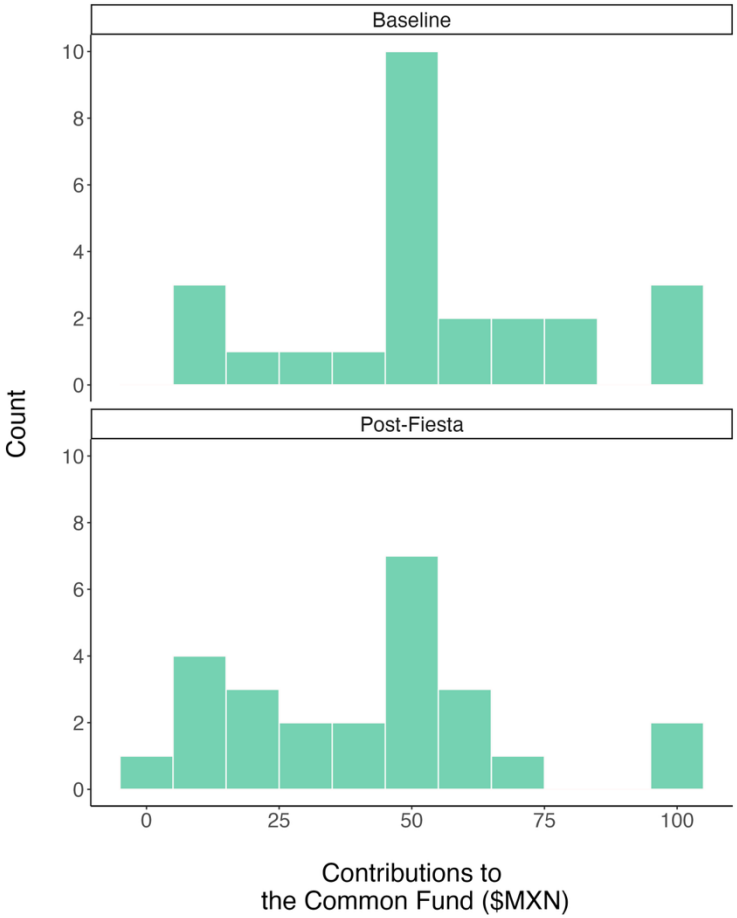

Figure S9. Histograms of Public Goods Game allocations to the Common Fund before and after the *fiesta.*

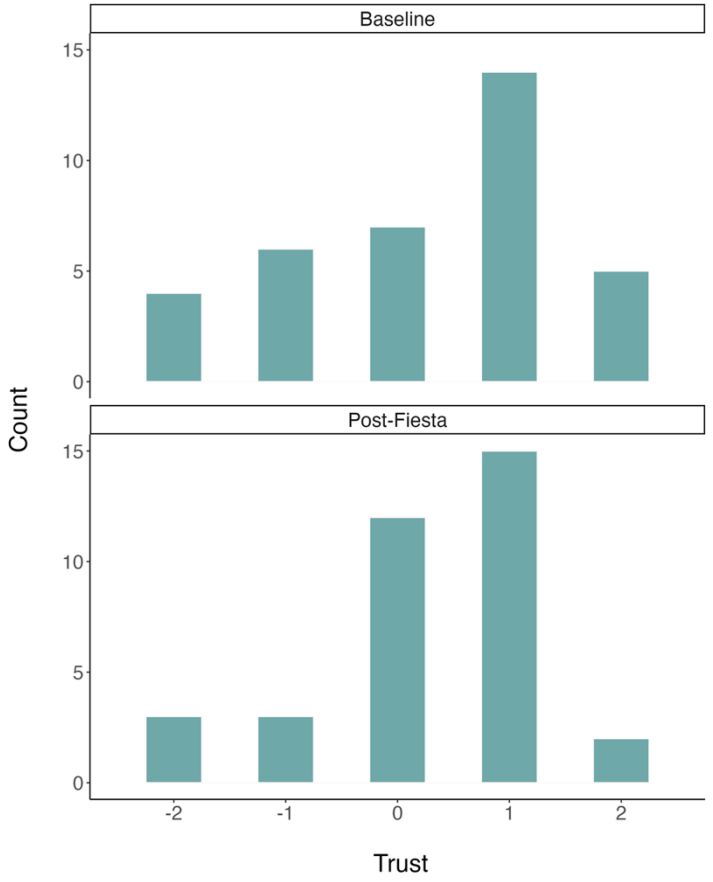

Figure S10. Histograms of responses to the trust question before and after the *fiesta.*

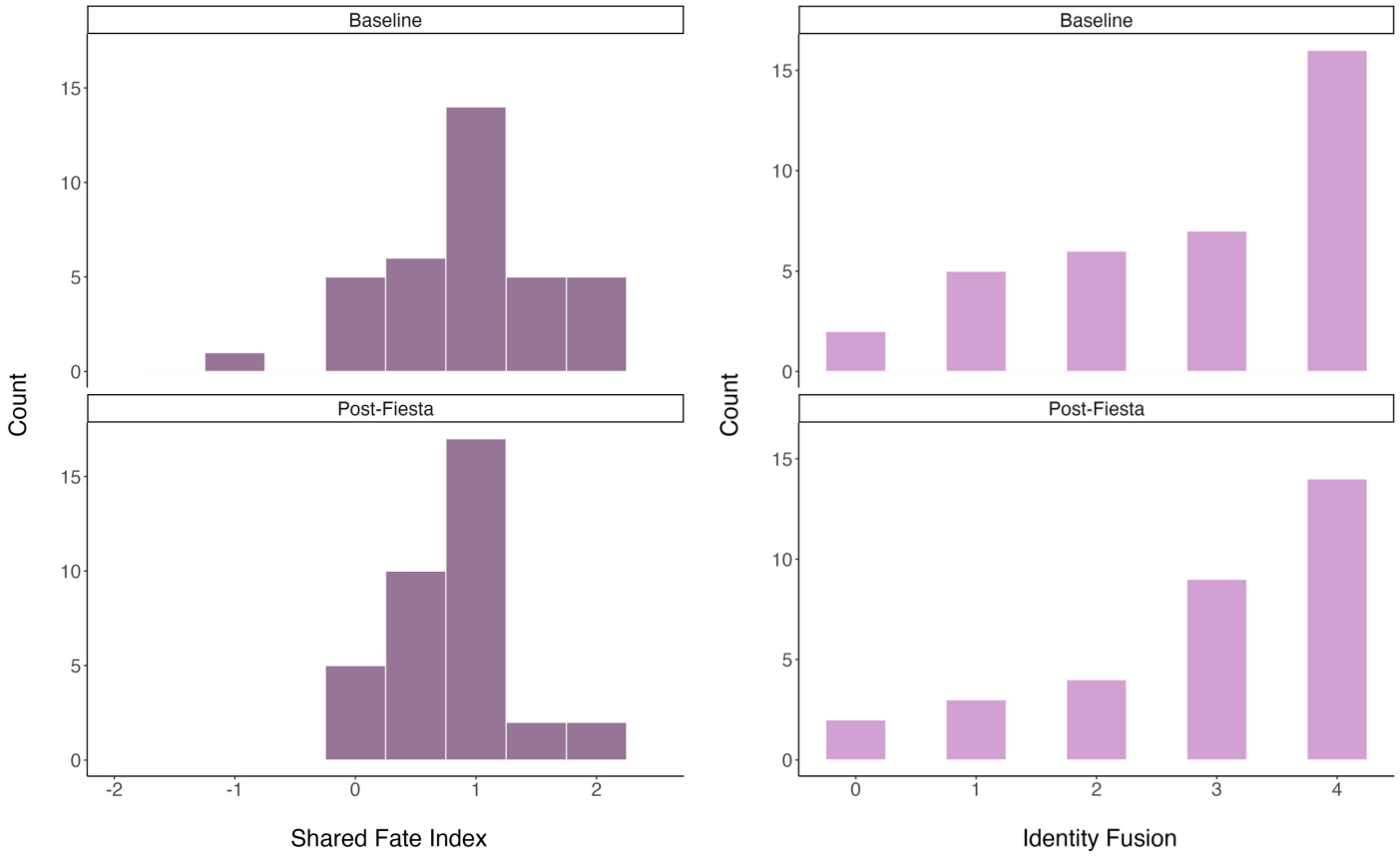

Figure S11. Histograms of survey responses to cohesion questions before and after the *fiesta.*

S4. Discussion

In seeking to understand the persistence of costly Oaxacan *fiestas*, a final possibility is that the annual patron saint *fiesta* is qualitatively different from smaller *fiestas* that communities celebrate throughout the year. Moreso than smaller *fiestas*, the patron saint *fiesta* is a time when migrants return to the village. Despite their physical absence, migrants still play an important role in community function. For example, they must pay annual taxes (*cooperaciones*) and men can be called back to serve *cargos*. The *fiesta* may help bind migrants to the community, ensuring their cooperation within the *usos y costumbres* system.

Table S1. Evidence of declines in ingroup altruism from baseline to post-*fiesta* (Full sample)

	Willingness to Help		*Donate to sick child*		*Bring food to widow*		*Stacked vignette measure*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Fiesta	–0.44* (–0.85, –0.02)	–0.74** (–1.20, –0.29)	–0.25 (–0.70, 0.21)	–0.22 (–0.65, 0.20)	–0.29 (–0.65, 0.07)	–0.28 (–0.68, 0.13)	–0.26 (–0.55, 0.03)	–0.25 (–0.55, 0.05)
Material security		0.23 (–0.07, 0.52)		0.17 (–0.08, 0.43)		–0.01 (–0.29, 0.27)		–0.10 (–0.11, 0.31)
Random Eff.								
Sub.Std. Dev.	0.47	0.76	0.30	0.60	0.65	0.75	0.42	0.58
Var. Std. Dev.							0.17	0.14
Obs*(N)*	72(36)	60(36)	72(36)	60(36)	72(36)	60(36)	144(36)	120(36)

Notes: LMER estimates with 95% confidence intervals. *Willingness to help* asked participants to state their level of agreement with the statement, "I would help a fellow community-member, even if they could never return the help." *Donate to a sick child* and *Bring food to widow* represent responses to vignettes wherein participants were asked how likely they would be to help a community member in need. All continuous variables have been standardized. Columns 7 & 8 show analysis of a stacked dataset, wherein the two vignette measures of altruism are collapsed into a single column (4 observations per participant). All models include random intercepts for participants, and the stacked models include crossed random intercepts for participants and altruism measure. Sample sizes show the number of observations, with the number of participants in parentheses.

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table S2. Evidence of declines in ingroup altruism from baseline to post-*fiesta* (Non-Evangelical sub-sample)

	Willingness to Help		*Donate to sick child*		*Bring food to widow*		*Stacked vignette measure*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Fiesta	–0.58* (–1.06, –0.10)	–0.74** (–1.26, –0.23)	–0.38 (–0.89, 0.14)	–0.21 (–0.63, 0.21)	–0.26 (–0.68, 0.16)	–0.24 (–0.69, 0.22)	–0.32 (–0.64, 0.01)	–0.22 (–0.52, 0.09)
Material security		0.12 (–0.17, 0.42)		0.07 (–0.16, 0.30)		–0.02 (–0.33, 0.29)		–0.01 (–0.18, 0.21)
Random Eff.								
Sub.Std. Dev.	0.34	0.59	0.00	0.33	0.72	0.79	0.38	0.43
Var. Std. Dev.							0.16	0.11
Obs*(N)*	54(27)	48(27)	54(27)	48(27)	54(27)	48(27)	108(27)	96(27)

Notes: LMER estimates with 95% confidence intervals. *Willingness to help* asked participants to state their level of agreement with the statement, "I would help a fellow community-member, even if they could never return the help." *Donate to a sick child* and *Bring food to widow* represent responses to vignettes wherein participants were asked how likely they would be to help a community member in need. All continuous variables have been standardized. Columns 7 & 8 show analysis of a stacked dataset, wherein the two vignette measures of altruism are collapsed into a single column (4 observations per participant). All models include random intercepts for participants, and the stacked models include crossed random intercepts for participants and altruism measure. Sample sizes show the number of observations, with the number of participants in parentheses.

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table S3. Altruism measures: Comparing main results and influence analysis results

	Willingness to Help		*Donate to sick child*		*Bring food to widow*		*Stacked vignette measures*	
	Main (1)	Influence (2)	Main (3)	Influence (4)	Main (5)	Influence (6)	Main (7)	Influence (8)
Fiesta	–0.44* (–0.85, –0.02)	–0.56*** (–0.87, –0.25)	–0.25 (–0.70, 0.21)	–0.22 (–0.52, 0.07)	–0.29 (–0.65, 0.07)	–0.48** (–0.80, –0.16)	–0.26 (–0.55, 0.03)	–0.25* (–0.48, –0.03)
Random Eff.								
Sub.Std. Dev.	0.47	0.50	0.30	0.36	0.65	0.75	0.42	0.43
Var. Std. Dev							0.17	0.11
Obs*(N)*	72(36)	66(33)	72(36)	68(34)	72(36)	66(33)	144(36)	136(34)

Notes: LMER estimates with 95% confidence intervals. *Willingness to help* asked participants to state their level of agreement with the statement, "I would help a fellow community-member, even if they could never return the help." *Donate to a sick child* and *Bring food to widow* represent responses to vignettes wherein participants were asked how likely they would be to help a community member in need. All continuous variables have been standardized. Columns titled "Main" show results using the full sample. Columns titled "Influence" show results from models where the most influential participants have been removed (DFBETA > 0.33, maximum of 10% of sample removed). All models include random intercepts for participants. Sample sizes show the number of observations, with the number of participants in parentheses.

*** p < 0.001, ** p <0.01, * p < 0.05

Table S4. Altruism measures: Comparing main results and influence analysis results (material security control)

	Willingness to Help		*Donate to sick child*		*Bring food to widow*		*Stacked vignette measures*	
	Main (1)	Influence (2)	Main (3)	Influence (4)	Main (5)	Influence (6)	Main (7)	Influence (8)
Fiesta	–0.74** (–1.20, –0.29)	–0.99*** (–1.37, –0.60)	–0.22 (–0.65, 0.20)	–0.36* (–0.67, –0.06)	–0.28 (–0.68, 0.13)	–0.39* –0.74, –0.04)	–0.25 (–0.55, 0.05)	–0.35* (–0.66, –0.03)
Material security	0.23 (–0.07, 0.52)	0.27 (-0.03, 0.58)	0.17 (–0.08, 0.43)	0.28* (0.03, 0.53)	–0.01 (–0.29, 0.27)	0.01 (–0.25, 0.28)	0.10 (–0.11, 0.31)	0.14 (–0.10, 0.38)
Random Eff.								
Sub.Std. Dev.	0.76	0.94	0.60	0.87	0.75	0.78	0.58	0.63
Var. Std. Dev							0.14	0.15
Obs*(N)*	60(36)	56(34)	60(36)	54(33)	60(36)	54(33)	120(36)	108(33)

Notes: LMER estimates with 95% confidence intervals. *Willingness to help* asked participants to state their level of agreement with the statement, "I would help a fellow community-member, even if they could never return the help." *Donate to a sick child* and *Bring food to widow* represent responses to vignettes wherein participants were asked how likely they would be to help a community member in need. All continuous variables have been standardized. Columns titled "Main" show results using the full sample. Columns titled "Influence" show results from models where the most influential participants have been removed (DFBETA > 0.33, maximum of 10% of sample removed). All models include random intercepts for participants. Sample sizes show the number of observations, with the number of participants in parentheses.

*** p < 0.001, ** p <0.01, * p < 0.05

Table S5. Public Goods Game: Comparing main results and influence analysis results

	Contributions to the Common Fund ($MXN)			
	Main	Main	Influence	Influence
	(1)	(2)	(3)	(4)
Fiesta	–12.00 (–25.48, 1.48)	–13.89+ (–28.34, 0.56)	–11.30+ (–22.84, 0.23)	–14.64* (–26.36, –2.92)
Material security		4.98 (–4.15, 14.11)		9.15* (0.32, 17.98)
Random Eff.				
Sub.Std. Dev.	12.17	11.92	15.35	17.24
Obs*(N)*	50(25)	48(25)	46(23)	44(23)

Notes: LMER estimates with 95% confidence intervals. *Contributions to the Common Fund* shows how many pesos of their $100MXN endowment participants contributed to the common fund in the Public Goods Game. *Material security* asked participants how certain they were that they would be able to "buy or produce enough food to eat in the coming month". This variable has been standardized, with higher values indicating greater certainty. Columns titled "Main" show results using the full sample. Columns titled "Influence" show results from models where the most influential participants have been removed (DFBETA > 0.4, maximum of 10% of sample removed). All models include random intercepts for participants. Sample sizes show the number of observations, with the number of participants in parentheses.

*** p < 0.001, ** p <0.01, * p < 0.05, + p < 0.06

References

Aktipis, A., Cronk, L., Alcock, J., Ayers, J. D., Baciu, C., Balliet, D., Boddy, A. M., Curry, O. S., Krems, J. A., Muñoz, A., Sullivan, D., Sznycer, D., Wilkinson, G. S., & Winfrey, P. (2018). Understanding cooperation through fitness interdependence. *Nature Human Behaviour*, *2*(7), 429–431. https://doi.org/10.1038/s41562-018-0378-4

Ayers, J. D., Sznycer, D., Sullivan, D., Beltrán, D. G., van den Akker, O. R., Muñoz, A. E., Hruschka, D. J., Cronk, L., & Aktipis, A. (2023). Fitness interdependence as indexed by shared fate: Factor structure and validity of a new measure. *Evolutionary Behavioral Sciences*, *17*(3), 259–284.

Curtin, C. M., Vasquez Martinez, N., Vásquez Vásquez, Y., & Henrich, J. (2026). *Zapotec cooperative institutions: The psychological and social mechanisms* [Unpublished manuscript].

Gómez, Á., Brooks, M. L., Buhrmester, M. D., Vázquez, A., Jetten, J., & Swann, 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), 918–933. https://doi.org/10.1037/a0022642

Hruschka, D. J., Munira, S., Jesmin, K., Hackman, J., & Tiokhin, L. (2018). Learning from failures of protocol in cross-cultural research. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(45), 11428–11434. https://doi.org/10.1073/pnas.1721166115

Instituto Nacional de Estadística Geographía e Informatica. (2020). *Censo de Población y Vivienda*. Instituto Nacional de Estadística, Geografía, e Informática. https://www.inegi.org.mx/programas/ccpv/2020/?ps=microdatos#Datos_abiertos

Páez, D., Rimé, B., Basabe, N., Wlodarczyk, A., & Zumeta, L. (2015). Psychosocial effects of perceived emotional synchrony in collective gatherings. *Journal of Personality and Social Psychology*, *108*(5), 711–729. https://doi.org/10.1037/pspi0000014

Purzycki, B. G., & Lang, M. (2019). Identity fusion, outgroup relations, and sacrifice: A cross-cultural test. *Cognition*, *186*(November 2018), 1–6. https://doi.org/10.1016/j.cognition.2019.01.015

Singh, P., Tewari, S., Kesberg, R., Karl, J. A., Bulbulia, J., & Fischer, R. (2020). Time investments in rituals are associated with social bonding, affect and subjective health: A longitudinal study of Diwali in two Indian communities: Diwali Time investment & Social Bonding. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1805). https://doi.org/10.1098/rstb.2019.0430

Swann, W. B., Gómez, A., Seyle, C. D., 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*, 995–1011.

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