Group-level ethnic composition influences altruistic punishment: public goods experiments among Gitanos and non-Gitanos in southern Spain

Espín, Antonio M. and Brañas-Garza, Pablo

Universidad de Granada, Universidad Loyola Andalucia, University of Nottingham

31 July 2019

Group-level ethnic composition influences altruistic punishment: public goods experiments among *Gitanos* and non-*Gitanos* in southern Spain

Antonio M. Espín[1*], Pablo Brañas-Garza[2], Juan F. Gamella[1], Benedikt Herrmann[3], Jesús Martín[1]

Authors' affiliation:
[1] Universidad de Granada, Spain.
[2] Universidad Loyola Andalucía, Spain.
[3] University of Nottingham, United Kingdom.
*Corresponding author: kanton@ugr.es

Abstract

Humans often punish non-cooperators in one-shot interactions among genetically-unrelated individuals. This so-called altruistic punishment poses an evolutionary puzzle because it enforces cooperation norms that benefit the whole group, but is costly for the punisher. Under the "big mistake" (or "mismatch") hypothesis, social behaviors such as punishment evolved by individual selection at a time when repeated interactions with kin prevailed, and modern humans "mistakenly" apply it in one-shot interactions with non-kin. In contrast, cultural group selection (CGS) theories emphasize cultural differences in normative behavior and the role of intergroup competition and punishment for the emergence of large-scale cooperation in the absence of genetic relatedness. We conducted a series of multilateral-cooperation economic experiments with a sample of Spanish Romani people (*Gitanos* or *Calé*), who represent a unique cultural group to analyze the nature of punishment: *Gitano* communities rely heavily on close kin-based networks, maintain high consanguinity rates, and display a particularly strong sense of ethnic identity. We observe that *Gitano* non-cooperators were not punished by co-ethnics in only-*Gitano* (ethnically) homogeneous groups but were harshly punished by both *Gitano* and non-*Gitano* males in mixed groups. Although largely consistent with CGS-related theories, these results can help better qualify some of their interpretations and predictions.

Keywords: cooperation, punishment, Gypsy/Roma, ethnicity, culture, evolution

Introduction

Humans possess an extraordinary capacity for large-scale cooperation and this stands as a theoretical puzzle across the biological and behavioral sciences. Mechanisms such as kin selection and direct and indirect reciprocity have been proposed as explanations for the evolution of cooperation in relatively small populations (Hamilton 1964; Nowak 2006). To explain prosocial behavior in large modern societies, however, kinship or reciprocity mechanisms seem to be insufficient[1] because cooperation is observed in ephemeral encounters among unrelated individuals; for instance, in voting, driving, paying taxes, recycling, market interactions, and warfare (Boyd and Richerson 1988; Gintis et al. 2003). Decentralized (peer) punishment of free-riders has been shown to be a crucial element for understanding the emergence of cooperation beyond kinship and small-scale groups (Boyd et al. 2003; Gintis et al. 2003; Henrich 2004; Sigmund 2007). So-called altruistic punishment is a kind of costly norm enforcement mechanism which cannot be explained by reputation or other traditional forms of reciprocity. Punishment is considered altruistic (in the biological sense) when the absolute benefits triggered by the enforcement of the cooperative norm are received by individuals other than the punisher (Fehr and Gächter 2002).

Even if groups in which peer punishment is allowed can outcompete those in which it is not due to the discouragement of free-riding (Gächter, Renner, and Sefton 2008; Sääksvuori, Mappes, and Puurtinen 2011; but see Herrmann, Thöni, and Gächter 2008), altruistic punishers are condemned to a lower evolutionary success within their group (Dreber et al. 2008). It turns out that the provision of a sanctioning system to prevent free-riding can be considered as a second-order social dilemma where individual and collective interests are in

[1] Note that indirect reciprocity (e.g., a reputation system) has been proposed as a mechanism to explain cooperation in sizeable groups. However, it is not hard to see how maintaining a system in which individual reputation is known by every other individual entails cognitive and monitoring costs which greatly increase with group size, thus limiting the effectiveness of indirect reciprocity to explain large-scale cooperation (Henrich and Muthukrishna 2021). This is consistent with results from evolutionary models (Panchanathan and Boyd 2003).

conflict (Fehr and Gächter 2002). Nevertheless, altruistic punishment is frequently observed in controlled experiments with unrelated human subjects, even in one-shot anonymous interactions (Espín et al. 2012; Fehr and Gächter 2002; Gächter and Herrmann 2009; Henrich et al. 2006). In fact, the neurobiological evidence suggests that people suffer disutility from observing uncooperative behaviors and derive pleasure from punishing wrongdoers (Crockett et al. 2013; de Quervain et al. 2004; Tabibnia and Lieberman 2007), which facilitates punishment decisions, even if they are costly. Yet the evolutionary basis of punishment behavior and its psychological underpinnings is subject to debate. Why do people pay irrecoverable costs to punish others?

The "big mistake" (or "mismatch") hypothesis (Cosmides and Tooby 1992; Delton et al. 2011; Delton and Krasnow 2017; Hagen and Hammerstein 2006; Krasnow et al. 2012; Lehmann et al. 2007; Tooby and Cosmides 1990) holds that the psychological mechanisms underlying group-beneficial behaviors, such as altruistic punishment, evolved in a period of human history in which nearly all social interactions were repeated and took place among close relatives. Thus, "traditional" reciprocity and/or kin selection mechanisms would lie behind the evolution of punishment, which emerged because under those circumstances punishing others benefits the individual's (direct or indirect) inclusive fitness, for instance, by reducing future exploitation by others. Such pan-human social psychology, so the argument goes, misfires in the behavior of modern humans, who "mistakenly" use altruistic punishment even in one-shot interactions with unrelated individuals (i.e., where it is no longer adaptive or fitness enhancing). It is argued, therefore, that human social psychology is programmed to differentiate between acquaintances and strangers only due to a desire to cultivate and maintain individually profitable, coalitional social-exchange relationships. Thus, the key elements to explain social behavior according to this line of argument are (coalitional) closeness and genetic relatedness. Different ecologies or environmental cues, however, would

lead to different expressions of the common evolved psychology and thus create behavioral variation. For the sake of conciseness, we will refer to the above constellation of ultimate and proximate explanations within the big-mistake tradition as "mismatch-related" theories.

On the other hand, following *cultural group selection* (CGS) theories and their associated "norm-psychology" perspective (Boyd et al. 2003; Chudek and Henrich 2011; Gintis et al. 2003; Henrich 2004, 2015; Henrich et al. 2010; Henrich and Muthukrishna 2021; Mesoudi 2016; Richerson et al. 2016; Richerson and Boyd 2008; Soltis, Boyd, and Richerson 1995), those proximate mechanisms (i.e., the negative emotions associated with the observation of uncooperative acts and the positive emotions associated with their punishment) may be particularly suited for solving the second-order dilemma of punishment—and hence the first-order dilemma of cooperation—in modern large-scale societies where one-shot interactions with non-relatives are common. Altruistic punishment would thus have been shaped following a complex process in which genes and culture co-evolve, with cultural adaptation being much more rapid than genetic adaptation. Under this account, different cultural groups develop the human "norm-psychology" (Chudek and Henrich 2011) differently in competition with other cultural groups. In particular, specific social behaviors which are advantageous for the group during intergroup competition are transmitted across individuals through social learning mechanisms (i.e., payoff- or frequency-biased imitation) and enforced through sanctions. Cultural groups with more group-beneficial norms, and consequently such norms themselves, are more likely to proliferate.[2] Behavioral variation would not be the result of current ecology alone, as implied by mismatch-related theories, but

[2] Intergroup competition in cultural evolution is a necessary condition for prosocial norms, rather than any other norms or behaviors, to evolve under CGS. This is due to the existence of multiple evolutionarily stable equilibria, with only some equilibria being "prosocial". Punishment can indeed stabilize any behavior (Boyd and Richerson 1992). However, intergroup competition does not need to imply extinction, or even intergroup violent conflict (although it is key in human history; Bowles 2006; Choi and Bowles 2007); there are other forms of intergroup competition such as prestige-biased group transmission, differential reproduction, and migration, or differential group survival/adaptation to ecological conditions without conflict (Henrich and Muthukrishna 2021; Richerson et al. 2016).

of its interaction with cultural history as well (thus, not denying evolutionary mismatch). Therefore, it is likely that some cultural groups use decentralized punishment of free-riding extensively, particularly those in which other mechanisms such as kinship and reputation are insufficient, while others are more lenient or most probably use it to enforce different norms. Under this account, human social psychology should be essentially programmed to differentiate between acquaintances and strangers and, more specifically, between ingroup and outgroup individuals, as this distinction is key to the success of one's own cultural group. Compared to mismatch-related theories, the role of relatedness and closeness is not as central in this line of thought, which instead emphasizes intergroup competition processes in cultural evolution. The collection of ultimate and proximate explanations within this tradition will be referred to as "CGS-related" theories.

To shed light on the nature of altruistic punishment, we conducted a series of lab-in-the-field economic experiments with a unique sample of Spanish Romani people (*Gitanos*, also referred to as *Calé*). Romani groups represent the largest ethnocultural minority in Europe.[3] Our experiments were primarily designed to explore how the ethnic composition of a cooperating group influences peer punishment. To do so, we brought together non-*Gitano* and *Gitano* people who, due to the cultural characteristics outlined below, could help answer important questions about the underpinnings of punishment behavior.

Gitanos constitute a paradigmatic case study for the purposes of this paper because: (i) kinship is at the core of their social life and organization even if they live a "modern" life, which in many other aspects resembles that of their non-*Gitano* neighbors (i.e., the majority Spanish population). Indeed, consanguinity rates within *Gitano* communities in the geographic area of the study are among the highest ever reported in Europe (Gamella 2020),

[3]Nonetheless, they have received little attention in experimental research. We are aware of only two studies analyzing the behavior of Romani people: Brañas-Garza, Cobo-Reyes, and Domínguez (2006) using the ultimatum game and Martín et al. (2019) using time discounting tasks. Behavior towards Romani people, but not their own behavior, is studied in Bauer et al. (2018).

at the upper bound of the range observed in traditional small-scale societies of hunter-gatherers and horticulturalists, which are considered to resemble the living conditions of ancestral humans (see Text S1 in the Supplementary Materials). *Gitanos* therefore constitute an exceptionally "rare" case. (ii) *Gitanos* display a strong sense of ethnic identity although in many ways share a bicultural identity (Benet-Martínez et al. 2002). While they mostly speak the majorities' languages and have adopted the religion and even a number of their neighbors' mores, they also maintain a strong and vibrant sense of themselves as a separate people. *Gitanos* try to preserve a separate ethnic identity, often reinventing their processes of differentiation, which are mainly based on reproductive strategies where specific factors including marriage, gender, and kin systems are crucial (Gamella and Martín 2007; Gay Blasco 1999; Martin and Gamella 2005). As a consequence, for example, even though *Gitanos* and non-*Gitanos* have cohabited the study area for more than 15 generations, mixed marriages have been traditionally rare (less than 5% for over two centuries in the study area). Although this is changing in areas where the integration of *Gitanos* in education and labor has been notable, according to our data ethnic intermarriage in the localities studied has not surpassed the 10% mark until very recently (Gamella and Álvarez-Roldán 2021). *Gitanos* and other European Romani groups (but not all) may constitute exceptional examples of ethnic resistance and integration at the same time. Interestingly, recent advances suggest that societies with more intensive kin-based institutions tend to display a stronger ethnocultural (group) identity, ingroup-outgroup differentiation, and ingroup loyalty (Henrich 2020; Schulz et al. 2019).

Gitanos, as other Romani groups, have developed a series of autonomous law-making processes that are often encoded in open-ended codes of norms, the *Gitano* Law. Although somewhat less elaborated than in Eastern European Romani groups (Marushiakova and Popov 2007; Weyrauch 2001), these processes are important in the effort to limit the escalation of

conflicts between families and kin networks, where the duty to defend and support family members is a central concern (San Román 2010).

There also exist fundamental differences in gender roles between *Gitanos* and non-*Gitanos*. In particular, although this is also changing in recent years, the most relevant difference for the focus of this paper is that *Gitano* norms prescribe women to assume a secondary role in the presence of males in public encounters (Gamella 2000; Gamella and Martín 2007; Gay Blasco 1999; San Román 2010), whereas normative principles of this type are not observed among non-*Gitanos*. See Text S1 in the Supplementary Materials for more ethnographic details.

Basic Design and Hypotheses

Before putting forward our research questions and hypotheses, we summarize the basic elements of the experimental design. We conducted our experiments with a total of 320 participants (mean age = 42.80 ± 18.42 SD, 59% females). We recruited *Gitano* and non-*Gitano* "ordinary people" from five small semi-rural towns with a large *Gitano* population in southern Spain. Participants played a one-shot public goods game with peer punishment (PGP) involving real monetary stakes in anonymous four-person groups. Given that participants only played one round of the game and groups were formed anonymously, no strategic concerns (e.g., about potential consequences in future interactions) were present for punishment decisions.

The experimental design comprises two conditions (between-subjects): participants played the PGP in either (i) *homogeneous* groups composed of either only *Gitanos* or only non-*Gitanos* or (ii) *mixed* groups with two *Gitano* and two non-*Gitano* members. Importantly, the two conditions were run in different sessions. Thus, ethnic identity was made particularly salient

in the mixed sessions because in the homogenous sessions there were only members of one's own cultural group. Ethnicity itself is often rather meaningless until the presence of "others" makes it relevant for social interaction and cultural identification processes (Brewer 1999; Dovidio, Gaertner, and Saguy 2008; Tajfel 1974). While among minority status groups, such as *Gitanos*, group identity is typically carried to every public environment (Pinel 1999), in the mixed condition the behavior of the two cultural groups could be directly compared by the participants, which should enhance the salience of intergroup encounter cues and hence of ethnic identity. Still, *Gitanos* and non-*Gitanos*, as minority and majority status groups, experience this ethnocultural difference asymmetrically (with asymmetries of power, position, and perspective, as well as more subtle experiential and interactional asymmetries; Brubaker et al. 2018).

Following standard procedures (Gächter and Herrmann 2009), participants in the PGP first made their cooperation decisions by means of (anonymously) allocating money from their €10 endowment to a group pot (any amount between €0 and €10). Contributions were doubled and evenly shared among the four group members. Thus, the more one contributes to the group pot (i.e., the public good), the larger the total group benefit, but the lower the decision maker's personal benefit, all else equal. This creates the classical social dilemma between collective and individual interests.

After all the participants had made their decisions, they were shown the contributions of each of the other three group members and allowed to spend part of their earnings in order to reduce others' earnings (punishment stage): €1 spent on punishment reduced the target individual's earnings by €3. It is important to remark that participants contributed knowing beforehand that they could be punished, which introduces strategic incentives to cooperate in order to avoid being punished. The reasons underlying contribution decisions are multiple and, therefore, cooperation does not constitute the main focus of our study.

Finally, by means of a subtle procedure which preserved participants' anonymity, we allowed participants in the mixed groups to match the ethnicity and contributions of each of the other three group members. Hence, our procedure let participants condition their punishment decisions on the target's ethnic identity. Note that this was not relevant in the homogenous groups since all four members were of the same cultural group. See Materials and Methods for a more detailed description of the experimental procedures.

Research questions and hypotheses

While the current study certainly has elements of an exploratory character, our research questions allow us to test a series of hypotheses about our participants' punishment behavior based on past empirical and theoretical evidence and our own ethnographic work (see Text S1 in the Supplementary Materials). To build these hypotheses, the main variables we consider are ethnicity, treatment condition, and gender. We provide a sketch of the arguments supporting each hypothesis here and extend on them in the Discussion section. Whenever possible, we compare predictions built upon CGS-related arguments with predictions emanating from mismatch-related theories. As mentioned, we test the norm-psychology account inherent to CGS by highlighting key differential cultural norms of *Gitanos* and non-*Gitanos* observed in our ethnographic work in the study area. This account states that human social psychology is unique in the animal kingdom because the human brain has differentially evolved to be highly sensitive to social norms, defined as "learned behavioral standards shared and enforced by a community" (Chudek and Henrich 2011:218). If the norm-psychology hypothesis is correct, *Gitanos*' and non-*Gitanos*' behavior in the experiment should reflect such differences in cultural norms, which work as proximate-level behavioral explanations driven by the internalization of the group's norms. In Text S1 in the

Supplementary Materials we explore some of these cultural differences and the associated (proximate-level) hypotheses in more detail, in particular those related to norm enforcement institutions and gender roles.

1) Is altruistic punishment more, or less, frequent in ethnocultural groups in which individuals are more strongly related (whether due to kinship or to closer/frequent interactions)?

If current punishment behavior represents the misfiring of a pan-human psychology (that emerged in an ancestral past where kinship- and closeness-based interactions prevailed, but is common to all modern humans) as prescribed by mismatch-related theories, a radical interpretation would predict few differences between the ethnic groups, as opposed to the predictions of CGS-related theories. However, it can also be argued that one should expect misfiring to be more prominent among *Gitanos*. In other words, if punishment evolved because it yields direct or indirect inclusive fitness benefits to the punisher but is "mistakenly" used in the current scenario due to the existence of cues evoking the ancestral scenario, *Gitanos* might in general punish free-riders *more* than non-*Gitanos* due to their higher genetic relatedness and closer daily-life relationships; socioecological conditions that are more similar to those faced by ancestral humans.

Nevertheless, the role of genetic relatedness for punishment seems particularly complex. On the one hand, the (reasonably expected) higher cooperation between kin might help punishment to proliferate under multilevel/group selection because the evolutionary disadvantage of altruistic punishers relative to non-punishing cooperators is reduced as average cooperation increases, thus rendering costly punishment less necessary (Boyd et al. 2003). The problem with this argument is that kinship cannot just enter the model to increase cooperation; genetic relatedness may have other consequences as well. Indeed, on the other

hand, individual within-group selection pressures seem to act against peer punishment in environments of high relatedness due to the negative impact of punishment on the fitness of individuals who share genes with the punisher and, consequently, on the punisher's inclusive fitness (Gardner and West 2004). Previous studies suggest that this might indeed be the case (e.g., Henrich and Henrich 2014; Schulz et al. 2019).

CGS theorists have long argued that peer punishment, in contrast to other sanctioning mechanisms such as centralized punishment or ostracizing, should be more frequently observed in larger, more complex societies with more impersonal interactions. Past cross-cultural studies support this prediction (Henrich et al. 2010; Marlowe et al. 2008). The reason is that the costs associated with punishing (e.g., damaging long-term relationships or the possibility of counter-punishment against the punisher or her family; to which we can add the aforementioned direct negative impact on the punisher's inclusive fitness) are higher in more tightly knit communities (Henrich and Muthukrishna 2021). Therefore, according to this line of thought, we should find that *Gitanos* use altruistic punishment *less* than non-*Gitanos* in the experiments. Our own ethnographic data, interpreted from a norm-psychology perspective, also aligns nicely with this prediction insofar as the culture of individual liberty observed in the *Gitano* population (Gamella 2011; San Román 2010) suggests that it is not the individual's responsibility to sanction co-ethnics' sporadic instances of non-cooperation; solidarity, and forgiveness may be the intuitive reaction (Brañas-Garza, Cobo-Reyes, and Domínguez 2006; see Text S1 in the Supplementary Materials).

2. Does the ethnic composition of the group influence punishment behavior?

The above hypotheses about ethnic differences were developed considering each ethnocultural group separately and are therefore focused on individuals' behavior in homogeneous groups. Now we turn to the mixed groups.

Following mismatch-related theories, the presence of members of the other cultural group in mixed groups might lead to a reduction in aggregate punishment, especially by *Gitanos*, compared to the homogeneous groups. This reduction would go hand in hand with the associated reduction of cues of genetic relatedness and the diminished likelihood of establishing coalitional social-exchange relationships among the interactants. Put differently, it is the group composition that is relevant to build a reputation of "formidability" that can be individually beneficial (Sell, Tooby, and Cosmides 2009). Reputation-driven psychological modules, such as emotions of anger, should be more easily evoked the more the environmental cues suggest that future interaction with group members will be likely. Thus, we should expect that both groups, especially the *Gitanos*, will punish less frequently in the mixed than the homogeneous condition.

Recent advances within the mismatch-related line of thought argue, however, that altruistic punishment may function as a deterrence mechanism to protect not only from potential future harmful acts against the self but also against valued others—where "valued" means that the punisher's fitness is dependent upon them for genetic or coalitional reasons (e.g., Delton and Krasnow 2017, Krasnow et al. 2012). Applied to intergroup encounters, ingroups should be more valued than outgroups and punishment might thus be used to defend them from future poor treatment (actually, to defend the punisher's interests linked to the ingroups). Although this extended deterrence logic has typically been used to explain third-party punishment—while the PGP has components of both second- and third-party punishment (see Discussion section)—it can to some extent be translated to our design: participants should punish outgroup free-riders the most and ingroup free-riders in the mixed groups the least, whereas the punishment of ingroup free-riders in the homogeneous groups should fall somewhere in between. The rationale is that, in the mixed groups, ingroups are the victims of outgroups' wrongdoing, whereas (leaving the punisher apart) only outgroups are the victims of ingroups'

wrongdoing. Therefore, it is evident that it is the ingroup's interests that should be protected through punishment. In the homogeneous groups, ingroups are both victims and perpetrators so it is less clear who should be defended. Previous research using the third-party punishment game supports the existence of such behavioral patterns (Bernhard, Fischbacher, and Fehr 2006; Delton and Krasnow 2017; Goette, Huffman, and Meier 2006; Jordan, McAuliffe, and Warneken 2014; Schiller, Baumgartner, and Knoch 2014). These patterns are expected to be sharper among *Gitanos*, who have stronger interests in ingroups due to both genetic and coalitional reasons.

A somewhat radical interpretation of CGS-related theories suggests that ingroup wrongdoers should be punished more strongly than outgroup wrongdoers, as punishment behavior is argued to play a fundamental role in maintaining ingroup cohesiveness (Chudek and Henrich 2011; Henrich et al. 2010; Richerson et al. 2016). In parallel, given the importance of intergroup conflict and parochialism for this account (Bowles 2006; Choi and Bowles 2007; Henrich and Muthukrishna 2021), participants are expected to punish outgroup (vs. ingroup) cooperators more spitefully/antisocially (Brañas-Garza et al. 2014; Herrmann, Thöni, and Gächter 2008) as harming the outgroup helps one's own cultural group outcompete other groups. Therefore, according to this interpretation of CGS-related theories, we should observe relatively more altruistic punishment of ingroup (vs. outgroup) free-riders and more antisocial punishment of outgroup (vs. ingroup) cooperators in both ethnic groups (Rusch 2014). Note that mismatch-related theories are typically silent about the role of antisocial punishment.

Nonetheless, a more comprehensive reading of CGS-related theories suggests that, even if cooperation seems to be a human moral universal (Curry, Whitehouse, and Mullins 2019), each cultural group should use sanctions to enforce those social norms which are particularly beneficial for the group (see the most recent extended synthesis in Henrich and Muthukrishna 2021). The different historical trajectories of *Gitanos* and non-*Gitanos*, with the former

traditionally being a discriminated-against minority (in Spain and elsewhere; Martín et al. 2019), might thus be associated with different group-level functional needs and domain-specific cooperation (Curry, Whitehouse, and Mullins 2019) and hence result in peer punishment being used to enforce different social norms. The next research question tackles this issue.

3. Do individuals punish ingroup free-riders differently when there are outgroups in the group?

As is often the case with ethnic minorities, compared to the non-*Gitano* majority, *Gitano* people display stronger group identity and higher group entitativity (i.e., the group is perceived to be a unified, single agent by outgroups and, consequently, the behavior of the whole group is often automatically identified with the behavior of its individual members; Gil-White 2001; Hamilton, Sherman, and Lickel 1998). Previous evidence suggests that individuals from groups with higher entitativity are more prone to feelings of collective responsibility when the group identity of an ingroup wrongdoer is salient, as in our mixed condition (Kardos et al. 2019). That is, during intergroup encounters, members of a wrongdoer's group often react with feelings of shame and anger and may take actions to protect the ingroup reputation, such as sanctioning the ingroup wrongdoer, and this is stronger in groups with sharper identification (Likel, Schmader, and Spanovic 2007; Marques, Yzerbyt, and Leyens 1988; Marques, Abrams, and Serodio 2001). According to this argument, we should expect that *Gitanos* display a stronger sense of collective responsibility and, therefore, punish ingroup free-riders more harshly in mixed than in homogeneous groups, as ethnic identity is made salient in the former condition. This argument also entails that non-*Gitanos* should inflict harsher punishment to ingroup free-riders in the mixed than in the homogeneous condition, but the effect is expected to be weaker due to the lower identification and group entitativity.

From the perspective of CGS-related theories, this hypothesis might be interpreted under the argument that different social norms are key for different cultural groups: while for non-*Gitanos* (displaying the characteristics of large, impersonal societies) a norm of fair generalized cooperation is expected to be crucial (Henrich et al. 2010), for *Gitanos* the key norm might be to protect the ingroup reputation against identity threats (Akerlof and Kranton 2000) spurred by the group's high degree of entitativity.

Mismatch-related theories do not provide such clear predictions about the impact of our experimental conditions on the punishment of ingroups because in the PGP (in contrast to other settings such as the third-party punishment game; see the Discussion section), both ingroups and outgroups are simultaneously both victims and observers of any uncooperative act, regardless of whether the wrongdoer is an ingroup or an outgroup member (see Delton and Krasnow 2017). Yet, as mentioned, an extended deterrence argument would predict that ingroup wrongdoers might receive less punishment in the mixed than in the homogeneous groups because in the former there are outgroup victims whose welfare is less strongly associated with the punisher's interests. This should be more clearly observed among *Gitanos*, whose interests are relatively more dependent upon the ingroups' individual welfare.

4. Does gender moderate the effect of ethnicity or group-level ethnic composition on punishment?

The fact that males tend to gain leadership relative to females in intergroup encounters does not seem to be contested, regardless of whether the focus is on cultural evolution and group selection or on genetic evolution and individual/sexual selection (Mathew and Boyd 2011; McDonald, Navarrete, and Van Vugt 2012). Both mismatch- and CGS-related theories would therefore yield similar predictions: we may expect to observe that females punish comparatively less in the mixed than in the homogeneous condition relative to males.

Building upon the ethnographic evidence mentioned earlier, however, the norm-psychology approach inherent to CGS-related theories suggests that *Gitano* females might punish less than *Gitano* males (and less than non-*Gitano* males and females as well) in both experimental conditions because males are always present in the interacting groups. Under these circumstances, *Gitano* norms indicate that females should let males lead the public interaction and thus probably the responsibility to punish non-cooperators. Mismatch-related theories, on the other hand, do not provide predictions for the existence of culture-specific gender differences.

Materials and Methods

Five semi-rural towns in southern Spain (Granada, Andalusia) with comparable demographic characteristics hosted our experiments: Benalúa de Guadix, Darro, Deifontes, Iznalloz, and Pedro Martínez (see Figure 1a). As a call for participation, a €5 show-up fee and a drink and *tapa* at the end of the experiment were offered. Recruitment of non-*Gitano* participants was mainly done through the town halls (the activity was publicly announced as a study for the University of Granada and individuals informed the staff about their interest in participating, although some people just showed up to the experiment and were able to participate if there were available slots). The town halls however did not provide such a good means to contact *Gitanos* since they are typically less involved in towns' official collective activities, so we needed to encourage the participation of *Gitanos* using other methods. Although we asked the town halls staff to advertise the event among *Gitano* families, we also relied on our fieldwork knowledge of *Gitano* families to recruit local members.

Two of the main researchers (AME and JFG) announced the study in several *Gitano* households from different family lines and asked our acquaintances there to bring their

relatives and friends ("*su gente*", "their folks") to the experiments. While it is true that this makes a difference in the recruitment method for *Gitano* and non-*Gitano* participants, it is important to note that (i) being unable to fill the sessions with *Gitano* participants was the main risk to be avoided, (ii) many of the non-*Gitano* participants also "brought some of their folks" to the experiment, (iii) the same process was used for all the experimental sessions (see below) so that any treatment condition effect on behavior cannot be attributed to differences in the recruiting method, (iv) due to our group's long relationship with the *Gitanos* in this area, the people contacted by the researchers covered a fairly representative share of the *Gitano* population in each town, (v) there is a small number of *Gitano* family lines in each town due to the high relatedness of all the *Gitano* inhabitants, thus, "their folks" were not simply their close family unit but typically included their extended family as well (potentially also friends), and these families tend to be very large, (vi) the system used to assign colors to people (those coming together tended to receive scarves of the same color; see below) minimized the probability that two folks interacted in the experiment because only one person per color was assigned to each PGP group. In sum, given these features that reduced the impact of the recruitment method, we consider that it did not dramatically influence the results. Yet, self-selection and non-representativeness can still be an issue, as in most lab and field experiments (see Exadaktylos et al. 2013 for a thorough discussion).

In each location, we ran two experimental sessions in a between-subjects design: one ethnically homogeneous session (either all *Gitanos*, in two locations, or all non-*Gitanos*, in three locations) and one ethnically mixed session (same number of *Gitanos* and non-*Gitanos*; one session in each of the five locations) where ethnic identity was made salient. We ensured that subjects in one session did not learn the ethnic composition of the other session prior to participating. In each of the 10 sessions, 32 participants played the one-shot PGP in eight independent groups of four people. The participants were initially evenly assigned one out of

four colors using visible colored scarves. Colors were assigned similarly in both sessions, which induced color assignment to be dependent on ethnicity in the mixed sessions since two of the colors were assigned to *Gitanos* and the other two colors to non-*Gitanos*. This procedure was unknown to the participants and was done by giving scarves of identical color to participants who showed up together. Since *Gitanos* and non-*Gitanos* always arrived separately, the resulting assignment of colors to ethnic groups was nearly perfect (see below).

In the mixed sessions, we subtly induced the participants to realize the link between colors and ethnicities prior to playing the game (in the homogenous sessions we made the composition of colors public as well in order to allow for comparability between conditions): the eight participants of each color were placed together wearing their scarves and photographed by an assistant in front of the other participants. This feature of the design allowed the participants to associate cooperation decisions to ethnicities (i.e., colors) and condition their punishment decisions upon the ethnicity of the target in mixed groups. Data from post-experimental interviews indicate that most participants were able to associate ethnicities to scarf colors in the mixed sessions (even if socially desirable responding might have reduced their willingness to acknowledge this). See Figure 1b for a representation of the structure of the experiment.

For the statistical analyses, we excluded seven participants: two *Gitanos* because they participated in a homogeneous non-*Gitano* session (we learned their ethnicity ex-post) and five individuals from four different mixed sessions because their ethnicity did not match their scarf color (including them does not qualitatively affect the results). The final sample consisted of 143 *Gitanos* and 170 non-*Gitanos*.

The basic elements of the PGP design have been reported elsewhere (Espín et al. 2012). Each four-person PGP group was composed of one randomly selected person from each (scarf)

color. Beyond colors, group membership was unknown. After deciding how much to contribute to a public good from an endowment of €10 (marginal per capita return = 0.5; thus each contributed euro cost the individual 50 cents but increased the earnings of each of the other three group members by 50 cents), the participants received feedback on their group partners' contributions and earnings in a color-based fashion and could then anonymously reduce other group members' payoffs at a personal cost (cost-to-impact ratio of punishment = 1:3). Finally, the participants were also asked to state the level of punishment they expected from each group partner (no monetary incentives were used for the expectations task). Figures S1 and S2 in the Supplementary Materials display the contribution and punishment decision cards, respectively. Several examples of all stages of the PGP were displayed on a whiteboard to facilitate understanding of the game rules. The instructions were explained by the same researcher (PBG) in all the sessions.

After the PGP, the participants completed an unrelated task. At the end of the experiment, they were privately asked to answer a set of socio-demographic questions and received their payment. Mean earnings from the PGP were €13.34 ± €4.08 (SD).

Statistical analysis

All statistical analyses were conducted using Stata v. 13 (Stata Corp). We implemented OLS regressions for the analysis of contributions to the public good and multilevel generalized linear mixed model (GLMM) regressions for the punishment decisions (i.e., amount reduced through punishment) with random effects on the group, the decision-maker, and the target individual to account for the interdependence of data at these three levels. All the regression results are reported in Tables S1–S3 in the Supplementary Materials. In the main text, we report the coefficient, standard error (SE), and two-tailed *P*-value for each contrast obtained

from regressions in columns 1a–5a. The reported standard errors are always robust to heteroskedasticity. The main (binary) explanatory variables in the regressions are the decision-maker's ethnicity (*Gitano* vs. non-*Gitano*), the experimental condition (mixed vs. homogeneous), and the decision-maker's gender (male vs. female), as well as their interactions. For the analysis of punishment behavior in the mixed groups, we also included the target's ethnicity (*Gitano* vs. non-*Gitano*) as the main explanatory variable. Secondary explanatory variables included the difference between the decision-maker's and the target's contributions to the public good (i.e., punisher's minus target's) and the mean contribution of the other two group members. All regressions are repeated, in adjacent columns (1b–5b), with controls for the decision-maker's age (ranging from 16 to 82; mean for *Gitanos* = 34.56 ± 13.60 SD; mean for non-*Gitanos* = 49.97 ± 18.97 SD; the difference is significant, $P < .01$, t-test) and household income (ranging from to 0 to 9, corresponding to "0 euros/month" and "more than 5,000 euros/month" bins, respectively; mean for *Gitanos* = 1.944 ± 1.211 SD; mean for non-*Gitanos* = 3.195 ± 1.564 SD; the difference is significant, $P < .01$, Mann–Whitney test) as potential confounding factors (Martín et al. 2019).

Ethics statement

All participants provided consent prior to participation. Oral informed consent was obtained because literacy was not a requirement to participate due to the (expected) low educational level of many participants; only being able to read and write numbers was required to participate. This study was conducted in accordance with the Declaration of Helsinki for human research. All participants were treated anonymously by assigning them a numerical code in accordance with Spanish Law 15/1999 on Personal Data Protection. No association was made between their real names and the results. As is standard in socio-economic

experiments, no ethic concerns were involved other than preserving the participants' anonymity. This procedure was checked and approved by the Vice-Dean of Research at the School of Economics of the University of Granada, the institution hosting the experiment.

Results

Contributions to the public good. The results of the participants' cooperation as measured by their contributions to the public good are displayed in Figure 2. No main effect of ethnicity (coef of *gitano* = -0.361 ± 0.321 SE, $P = .26$; Table S1, column 1a) or condition (coef of *mixed* = 0.306 ± 0.312 SE, $P = .33$; Table S1, column 1a) on contributions was found. The interaction between these two variables was not significant either (coef of *gitano X mixed* = -0.733 ± 0.620 SE, $P = .24$; Table S1, column 2a) and all possible comparisons report $P > .10$ according to joint-significance Wald tests on the model estimates. Adding controls for age and household income does not qualitatively change the results (Table S1, columns 1b and 2b). Therefore, contributions did not differ between ethnic groups (in aggregate or within each condition) or between conditions (in aggregate or within each ethnic group). Contribution levels were relatively high (well above 60% of the endowment on average; see Ledyard 1995). Given that the threat of punishment introduces incentives to cooperate strategically and therefore contributions do not necessarily reflect a "pure" preference for cooperation, the finding of similar average contribution levels across cultural groups and conditions could be due to multiple factors. Note that, due to this multiplicity of motives, we did not put forward any hypotheses about the groups' cooperation levels and all the analyses on contribution behavior are thus exploratory.

However, we observed a significant interaction between gender and condition (coef of *mixed X male* = -1.543 ± 0.641 SE, $P = .02$; Table S1, column 4a; see Figure 2b and 2c). Across

both cultural groups (apparently more clearly among non-*Gitanos* although the three-way interaction ethnicity X condition X gender was not significant, coef = 0.613 ± 1.336 SE, $P = .65$; Table S1, column 5a), we found that females contributed more in the mixed than in the homogenous groups (coef of *mixed* = 0.921 ± 0.384 SE, $P = .02$; Wald test on Table S1, column 4a), while the opposite was observed for males (although not significantly so; coef of *mixed* = -0.622 ± 0.515 SE, $P = .23$; Wald test on Table S1, column 4a). As a result, males cooperated significantly less than females in the mixed groups (coef of *male* = -0.938 ± 0.435 SE, $P = .03$; Wald test on Table S1, column 4a but similarly in the homogeneous groups (coef of *male* = 0.605 ± 0.486 SE, $P = .21$; Wald test on Table S1, column 4a). Again, controlling for age and household income does not qualitatively affect the results (Table S1, columns 4b and 5b).

Aggregate punishment levels. Figure 3 summarizes the results regarding punishment behavior. We observed a significant main effect of ethnicity, indicating that, in general, *Gitanos* punished less than non-*Gitanos* (coef of *gitano* = -0.362 ± 0.116 SE, $P < .01$; Table S2, column 1a). The treatment condition did not yield a significant estimate (coef of *mixed* = -0.065 ± 0.148 SE, $P = .66$; Table S2, column 1a). A significant ethnicity X condition interaction (coef of *gitano X mixed* = 0.807 ± 0.228 SE, $P < .01$; Table S2, column 2a) reveals that *Gitanos* punished much less than their non-*Gitano* counterparts in the homogeneous groups (coef of *gitano* = -0.870 ± 0.156 SE, $P < 0.01$; Wald test on Table S2, column 2a), but there were no ethnic differences in the mixed groups (coef of *gitano* = -0.063 ± 0.157 SE, $P = .69$; Wald test on Table S2, column 2a; see Figure 3a). The intergroup encounter triggered by the mixed condition thus exerted substantial and differential effects on both sides: *Gitanos* increased their punishment level (coef of *mixed* = 0.389 ± 0.168 SE, $P = .02$; Wald test on Table S2, column 2a) while non-*Gitanos* reduced it (coef of *mixed* = -0.418 ± 0.193 SE,

P = .03; Wald test on Table S2, column 2a), as compared to the homogenous condition.

There was also a significant interaction between ethnicity and gender on punishment (coef of *gitano X male* = 0.574 ± 0.231 SE, *P* = .01; Table S2, column 3a). This stems from a higher level of punishment implemented by males compared to females among *Gitanos* (coef of *male* = 0.427 ± 0.130 SE, *P* < .01; Wald test on Table S2, column 3a). No aggregate gender difference in punishment was found among non-*Gitanos* (coef of *male* = -0.147 ± 0.189 SE, *P* = .44; Wald test on Table S2, column 3a).

Finally, a significant interaction was also found between condition and gender (coef of *mixed X male* = 1.084 ± 0.223 SE, *P* < .01; Table S2, column 4a). Specifically, we observed a higher level of punishment by males (coef of *mixed* = 0.593 ± 0.198 SE, *P* < .01; Wald test on Table S2, column 4a) and a lower level of punishment by females (coef of *mixed* = -0.492 ± 0.158 SE, *P* < .01; Wald test on Table S2, column 4a) in the mixed than the homogenous groups (see Figure 3b and 3c). This results in males punishing less than females in the homogenous groups (coef of *male* = -0.373 ± 0.136 SE, *P* < .01; Wald test on Table S2, column 4a) but more than females in the mixed groups (coef of *male* = 0.712 ± 0.174 SE, *P* < .01; Wald test on Table S2, column 4a). Although the three-way interaction ethnicity X condition X gender was not significant (coef = -0.476 ± 0.434 SE, *P* = .27; Wald test on Table S2, column 5a), it can be seen that *Gitano* females almost never used punishment in either condition. In other words, punishment by *Gitano* females was nearly inexistent regardless of the condition whereas the level of punishment implemented by *Gitano* males, which was negligible in the homogeneous groups, turned out to be rather high in the mixed groups. Among non-*Gitanos*, females punished less while males punished more in the mixed than in the homogeneous groups. As before, adding controls for age and household income does not alter any of the above findings (Table S2, columns 1b–5b).

Altruistic and antisocial punishment. In all the above regressions, the higher the difference between the punisher's contribution and the target's contribution (punisher's minus target's), the stronger the punishment (in all cases, coef of *differ* > 0.07, SE < 0.02, $P < .01$; Table S2, columns 1a–5a), thus indicating that more intense free-riding receives firmer punishment, as is standard in the literature (Espín et al. 2012; Fehr and Gächter 2002; Herrmann et al. 2008). However, we also observe some instances of spiteful, antisocial punishment targeted at cooperators. When disentangling between "altruistic" punishment (the target contributed less than the punisher) and "antisocial" punishment (the target contributed more than the punisher) in panels (a) and (b) of Figure 4, we see that the rather strong punishment implemented by *Gitanos*, in particular males (panels c and d break down the data by gender), in the mixed compared to the homogeneous groups is due almost uniquely to altruistic punishment since their level of antisocial punishment was still very low in the mixed groups. The remaining results mentioned above do not appear to crucially depend, at least qualitatively, on whether punishment is altruistic or antisocial.

Ethnocultural identities and punishment in mixed groups. It remains to be determined whether participants punished differently in the mixed groups depending on the cultural identity of the target (recall that the punisher knew the target's ethnicity but not her personal identity). In Figure 5, we display the mean punishment levels imposed on *Gitano* and non-*Gitano* targets in the mixed groups. We find that, *regardless of the punisher's ethnicity*, *Gitano* targets received less antisocial punishment and more altruistic punishment than non-*Gitano* targets for the same behaviors (significant interaction between target's ethnicity and contribution difference: coef of *targetgit X differ* = 0.105 ± 0.037 SE, $P < .01$; Table S3,

column 4a; the three-way interaction with punisher's ethnicity was not significant: coef = 0.057 ± 0.074 SE, $P = .45$; Table S3, column 5a; see Figure 5a and 5b). Put differently, both *Gitano* and non-*Gitano* punishers were more responsive to the distance between their own and the target's contribution (i.e., to the relative level of free-riding) when the target was *Gitano* than when the target was non-*Gitano*. *Gitano* targets got punished significantly less than non-*Gitano* targets when they cooperated more than the punisher (coef of *targetgit* between -0.323 ± 0.162 SE and -0.954 ± 0.347 SE, $P < .05$, for differences between €4 and €10; Wald test on Table S3, column 4a), whereas *Gitano* targets got punished more than non-*Gitano* ones when they cooperated less than the punisher (coef of *targetgit* between 0.414 ± 0.213 SE and 1.151 ± 0.439 SE, $P < .05$, for differences between €3 and €10; Wald test on Table S3, column 4a). Still, note that antisocial punishment was much less frequent than altruistic punishment. As can be seen in Figure 5c and 5d, the difference in altruistic punishment between *Gitano* and non-*Gitano* targets is due solely to male punishers, whereas the difference in antisocial punishment is similar across genders, although it appears to be stronger among non-*Gitano* female punishers. All these results also remain after controlling for age and household income (Table S3, columns 1b–5b).

To summarize, in contrast to the high punishment levels observed among non-*Gitanos*, *Gitanos* practically did not punish the misbehavior of other *Gitanos* in the homogeneous groups but (in particular males) severely punished such behavior in the mixed groups with non-*Gitanos*. Non-*Gitano* males, on the other hand, also retaliated more harshly against *Gitano* free-riders than against non-*Gitano* ones in the mixed groups. Regarding the antisocial punishment of cooperators, the results are somehow weaker: while participants, regardless of their ethnicity, tended to target more punishment at non-*Gitano* than *Gitano* cooperators in the mixed groups, the levels of antisocial punishment were relatively low (especially compared to those of altruistic punishment).

A closer look into the basic competing hypotheses. In Figure 6a and 6b we rearrange the above results regarding altruistic and antisocial punishment in a manner that facilitates comparison with the testable predictions outlined earlier. We employ the following notation: I = punishment targeted at ingroups; O = punishment targeted at outgroups. Subscripts G, NG, H, and M refer to *Gitano* punishers, non-*Gitano* punishers, homogeneous groups, and mixed groups, respectively.

Our research question (1) has a clear answer: with regards to altruistic punishment (Figure 6a), from the homogeneous condition we observe that $I_G < I_{NG}$ ($P < .01$). Thus, the punishment targeted at ingroup free-riders in the homogeneous groups is higher among non-*Gitanos* than among *Gitanos*, as predicted by the hypothesis built upon CGS-related theories.

Our research question (2) has a more complex answer, however. In the mixed groups, we can see that I > O in altruistic punishment (Figure 6a) holds for *Gitanos*, indicating that ingroup free-riders get punished more harshly than outgroup ones, whereas the opposite (I < O) is true for non-*Gitanos* ($P < .05$ for differences between the punisher's and the target's contributions larger than €3 in both cases; see above). With regards to the antisocial punishment of cooperators in the mixed groups (Figure 6b), we observe I < O for *Gitanos*, that is, outgroup cooperators get punished more harshly than ingroup ones, but the opposite (I > O) occurs for non-*Gitanos* ($P < .05$ for differences between the punisher's and the target's contributions larger than €4 in both cases; see above) Here, the results for *Gitano* punishers, but not for non-*Gitano* punishers, match the CGS-based predictions. The behavior of non-*Gitano* punishers is more in line with the predictions of mismatch-related theories to the extent that they reduce their level of altruistic punishment (remember that this account does not yield predictions about antisocial punishment) in the mixed compared to the homogeneous groups ($P = .03$). Moreover, although only in the case of males, non-*Gitanos* punish somehow according to the predictions of an extended deterrence account, that is, $O > I_H > I_M$: outgroup

free-riders get punished more than ingroup free-riders in the mixed groups, whereas ingroup free-riders in the homogenous groups receive intermediate levels of punishment (which links with our research question 3). Nonetheless, a mismatch-related approach would predict these patterns to be more evident among *Gitanos* than non-*Gitanos*.

The answer to our research question (3) is also intricate. Regarding altruistic punishment (Figure 6a), we observe that $I_H < I_M$ holds among *Gitanos* ($P < .01$ for all possible differences between the punisher's and the target's contributions), meaning that ingroup free-riders get punished more in the mixed than in the homogeneous groups, while the opposite is observed among non-*Gitanos*, albeit marginally ($P < .10$ for differences between the punisher's and the target's contributions larger than €1). Again, the CGS-based hypothesis gives a good approximation to the behavior of *Gitano* punishers but not to that of non-*Gitano* punishers, which is closer to the general predictions of mismatch-related models.

Our research question (4) tackles gender differences. Figure 6c displays the results on altruistic punishment for male and female punishers separately. As mentioned above, *Gitano* females practically did not punish in any condition. In addition, the $I_G < I_{NG}$ finding from the homogeneous groups and the $I_H > I_M$ finding for non-*Gitanos* hold qualitatively regardless of the punisher's gender, whereas the $I > O$ ($I < O$) observed among *Gitanos* (non-*Gitanos*) in the mixed groups as well as the $I_H < I_M$ among *Gitanos* are only driven by male punishers. Therefore, these gender differences are largely consistent with the "male warrior" hypothesis stating that males should punish comparatively more in the mixed vs. homogeneous groups relative to females. The norm-psychology prediction based on our ethnographic record that *Gitano* females should barely punish in either condition due to the presence of males is also supported by the data.

Breaking down the results on antisocial punishment by gender in Figure 6d, we see that the

I < O finding for *Gitano* punishers holds qualitatively for both males and females (with the disclaimer that females punish very little), whereas the I > O observed for non-*Gitanos* is only driven by male punishers. Note that these results on gender differences in antisocial punishment are shown for the sake of completeness even though we did not have any ex-ante prediction and are thus exploratory.

Discussion

Gitanos barely used punishment in the homogeneous groups, in contrast to non-*Gitanos* who used it rather frequently. These data seem to be inconsistent with a mismatch-related interpretation. Under this hypothesis, *Gitanos* should, on average, punish similarly or slightly more than non-*Gitanos* because their social organization is more heavily based on close family networks and relatedness. The explanation would be that group-beneficial behaviors such as altruistic punishment evolved at a time when nearly all social interactions were among relatives, and reputation was always at stake. Such an evolved psychology should be equally or more clearly displayed by *Gitanos* who are "still" heavily organized around kinship and close relationships compared to non-*Gitanos*, who tend to have more frequent sporadic encounters with non-relatives. Our results do not support this prediction.

On the other hand, if altruistic punishment is particularly important for the enforcement of cooperation among non-kin in large-scale societies, as argued by theorists of cultural group selection, non-*Gitanos* should punish more than *Gitanos* in homogeneous groups. This is what we observe. The results of Henrich and Henrich (2014) suggest that relatedness might reduce the willingness to punish others, since they found that individuals more genetically related to the average member of the "yavusa" in a Yasawan sample (Fiji Islands) tended to punish less

as third-party observers. Moreover, in such a highly genetically-related population, punishment was comparatively infrequent, and zero offers were very often accepted in both ultimatum and third-party punishment games, whereas actual offers were on average quite high (i.e., "fair"). This matches the ultimatum game results of Brañas-Garza, Cobo-Reyes, and Domínguez (2006) with a sample of Spanish *Gitanos* in Madrid, where high levels of cooperation were observed in the form of high offers even though much lower offers would have gone unpunished. In cultural groups organized around tight kinship-based networks, peer punishment may not be favored to enforce daily-life group cooperation if other mechanisms such as gossiping or centralized punishment institutions represent lower-cost solutions (given the short-run negative impact of punishment on the fitness of individuals who share genes with the punisher as well as the extra costs associated, e.g., to counter-punishment; Henrich and Muthukrishna 2021). Indeed, previous theoretical evidence suggests that punishment is typically selected against in environments of high genetic relatedness (Gardner and West 2004). Recent advances also indicate that public, multilateral cooperation might evolve by kin selection in sizeable groups in the absence of punishment if genetic relatedness is strong enough (values observed today in small-scale populations may suffice, e.g., Walker 2014) so that indirect inclusive fitness benefits act as a sufficiently powerful cooperation-enhancing force (Rusch 2018). In a related manner, experimental research suggests that cooperation, but *not* punishment, increases with cues of kin density in PGP groups (Krupp, Debruine, and Barclay 2008). The exact role of genetic relatedness (between the punisher, the victim(s), and the wrongdoer) for punishment behavior is yet to be systematically assessed, however; future research should provide such a systematic evaluation.

In addition, as opposed to the arguments of mismatch-related theories, the existence of cultural selection processes and a norm-psychology suggests that the same behavior (i.e., punishment of members of one's own cultural group) may be modulated by between-group

vs. within-group encounters in a culture-specific way. This is again what we observe.

Gitanos (but only males), who have a strong sense of ethnic identity, targeted punishment at *Gitano* wrongdoers when they interacted with non-*Gitanos* in the mixed groups but not in the only-*Gitano* homogeneous groups. At the proximate level, we interpret this result as reflecting that *Gitano* males use punishment only in response to a clear threat to group identity (Akerlof and Kranton 2000): that of being seen as less cooperative than non-*Gitanos*. The negative emotions triggering punishment (Crockett et al. 2013; Fehr and Gächter 2002) among *Gitanos* would thus emanate from the possibility of comparison between the two ethnic groups. Previous research indicates that, during intergroup contact, feelings of identity threat and collective responsibility are particularly likely to be aroused among individuals with a stronger group identification (Dovidio, Gaertner, and Saguy 2008; Marques et al. 1988; Hamilton et al. 1998; Kardos et al. 2019). It can thus be inferred that the key norm for *Gitanos* (that which is to be enforced through punishment) is not cooperation *per se*, but preserving an ethnic identity of which they are proud.[4] To a large extent, this result is coherent with previous results from ultimatum game experiments (McLeish and Oxoby 2007, 2011; Mendoza, Lane, and Amodio 2014) and multilateral gift-giving (non-standard) third-party punishment games (Shinada, Yamagishi, and Ohmura 2004) using identity manipulations.

However, the latter finding seems at odds with most results from standard third-party punishment experiments in which harsher punishment has been observed when the victim is an ingroup of the third-party (i.e., the punisher) and the norm violator is an outgroup,

[4] In fact, in the homogeneous condition, a common comment by *Gitano* participants during the post-experimental interview when informally asked about their perception of punishment opportunities (i.e., "the possibility of reducing others' earnings") was that punishing others makes no sense at all. "Destroying others' money and paying for it!" (subject #25) was seen as something weird, irrational, and very negative by *Gitanos* in the homogeneous condition. Comments of this type were inexistent in the mixed condition (as well as in the only-non-*Gitano* homogeneous condition), as if the reasons for punishing others were evident for everyone. In fact, even though the beliefs elicitation was not incentivized and should therefore be taken with caution, participants' expectations seem to match their behavior to a large extent: *Gitanos* expected much less punishment than non-*Gitanos* in the homogeneous groups ($P < .01$; same regression specification as for punishment decisions) and expected more punishment in the mixed than in the homogeneous groups ($P = .05$; Wald test).

compared to other combinations (Bernhard, Fischbacher, and Fehr 2006; Delton and Krasnow 2017; Goette, Huffman, and Meier 2006; Jordan, McAuliffe, and Warneken 2014; Schiller, Baumgartner, and Knoch 2014; but see Shinada, Yamagishi, and Ohmura (2004) for a non-standard design with different results). In contrast to results with adults, a recent third-party punishment experiment found that three to four-year-olds, but not older children (see also Jordan, McAuliffe, and Warneken 2014), inflict harsher punishment on ingroup than outgroup norm-violators (Yudkin, Van Bavel, and Rhodes 2019). Yet there are important differences between the multilateral cooperation environment of our PGP and the framework posed by the third-party punishment game in those experiments. For instance, the punisher is directly affected by the norm violation in the former but not in the latter. Also, both ingroups and outgroups can be victims (and observers) of the norm violation *at the same time* in the PGP but not in the third-party punishment game (see Delton and Krasnow 2017). Likewise, punishers might have been more cooperative than the target, or less, in the third-party punishment game. However, this fundamental detail—which informs about the true (altruistic vs. antisocial) nature of punishment (Brañas-Garza et al. 2014; Espín et al. 2015; Herrmann, Thöni, and Gächter 2008)—is by design unknown (but see Shinada, Yamagishi, and Ohmura 2004), in contrast to the PGP. Exploring the possible reasons for the inconsistencies between experimental frameworks (and between young children and adult behavior) is an interesting endeavor for future research.

Non-*Gitano* males' sanctioning behavior, on the other hand, seems closer to what previous experiments using (standard) third-party punishment games with adults have shown: they punish outgroup wrongdoers harshly but not ingroup ones in mixed groups. Indeed, the lowest level of altruistic punishment by non-*Gitano* males is observed when the wrongdoer is an ingroup and there are outgroup "third-party" victims, whereas the maximum level of punishment is targeted at outgroup wrongdoers when there are ingroup third-party victims.

When both the wrongdoer and the third-party victims are ingroups (i.e., in homogenous groups), their punishment remains at intermediate levels. Seen in this way, these behavioral patterns resemble previous observations from third-party punishment games with adults (see, for instance, Bernhard, Fischbacher, and Fehr 2006). Non-*Gitanos*' punishment behavior in mixed groups, therefore, seems inconsistent with the basic predictions of CGS-related theories. These patterns are instead more coherent with an extended deterrence explanation of punishment behavior (i.e., that punishment serves to deter future poor treatment of the self and valued others), as put forward by advocates of the mismatch hypothesis (e.g., Delton and Krasnow 2017). Although such an approach would suggest that these patterns should be more evident among *Gitanos* than non-*Gitanos* due to the stronger genetic and coalitional links with ingroups, we observe the opposite. Ultimately, the current results indicate that a deterrence function of punishment can be erroneously inferred if experiments are conducted with the "wrong" population or without comparing cultural groups of different status. It might be that the majority status of non-*Gitanos* and the associated lower strength of group identity and entitativity (see below) contribute to explaining this finding and the discrepancy with *Gitanos*' punishment behavior, which aligns well with the predictions of CGS-related theories in both the homogeneous and mixed groups.

In addition, we find some indication that *Gitanos* spitefully punished non-*Gitano* cooperators more than *Gitano* ones (i.e., more antisocial punishment targeted at outgroups than ingroups). This result is in line with the parochialism prediction of CGS-related theories as well (e.g., Choi and Bowles 2007), but the level of antisocial punishment in the mixed groups was perhaps too low to draw any firm conclusion.

Taken together, these results highlight the complexity of inter-ethnic relationships for both the provision of public goods and the enforcement of cooperation, as well as for their experimental analysis (see, e.g., Gil-White 2004). The role of majority versus minority status

of groups (and inherently the associated varying degrees of group entitativity; Hamilton et al. 1998), which has been largely overlooked in previous research on punishment behavior in intergroup encounters using economic games, might be crucial. As mentioned, *Gitano* non-cooperators were firmly punished by other (male) *Gitanos* in mixed groups, but also by (male) non-*Gitanos*. The fact that ethnic minorities, and Romani groups in particular, are often perceived as if not following the collective action norms of the majority (Bauer et al. 2018; Martin and Gamella 2005; Marushiakova and Popov 2007; Weyrauch 2001) and as potentially violent in their reactions to the majority's enforcement institutions (Gay Blasco 1999; San Román 2010) may explain the strong punishment of *Gitano* wrongdoers by non-*Gitano* males. This result could be reflecting the opportunity provided by the anonymous experimental setting for the majority to sanction the minority without fearing retaliation and is probably symptomatic of a sense of moral superiority (Brewer 1999) or pretended assimilation (Dovidio, Gaertner, and Saguy 2008). Further research should explore these possibilities in greater depth. Note that non-*Gitanos* typically do not share such a strong group identity and entitativity as *Gitanos* due, in part, to their majority status. Indeed, groups' majority/minority status is a predictable, although imperfect, correlate of group identity strength that shapes intergroup encounters in many ways (Dovidio, Gaertner, and Saguy 2008). Previous evidence indicates that members of majority status groups are typically more concerned with not being perceived as prejudiced by the minority, whereas members of minority groups are concerned with becoming the target of the majority's prejudice (Shelton 2003; Tropp and Pettigrew 2005). Since an extended stereotype is that Romani people do not contribute to the commons and display low compliance with the majority collective action norms (Bauer et al. 2018), following those arguments, it might be natural that both non-*Gitanos* and *Gitanos*, although for different reasons, punish acts that confirm the stereotype (i.e., *Gitanos* not cooperating) more firmly than acts that contradict it (i.e., non-*Gitanos* free-riding or *Gitanos* cooperating). This would be consistent with our findings.

An important aspect uncovered by our experiments relates to the impact of gender roles within as well as across cultural groups. While females contribute more in mixed than homogeneous groups, the opposite is observed for males. Also, in contrast to what we see among females, males punish generally more in mixed than homogeneous groups (consistent with our hypothesis based on a "male-warrior" account; Mathew and Boyd 2011; McDonald, Navarrete, and Van Vugt 2012). These two results hold similarly for both *Gitano* and non-*Gitano* participants, thus suggesting the existence of gender differences common to both cultural groups. One candidate proximate force underlying such gender differences in contributions and punishment is risk aversion (recall that we did not have predictions about contributions but only about punishment). If mixed groups are perceived as risky environments due to the presence of outgroups, probably the safest strategy is to avoid conflict by cooperating and not punishing others (to the extent that the punished individual cannot learn the ethnic identity of the punisher, punishment not only of outgroups but also of ingroups may trigger conflict). Since there is abundant evidence that, at least in patriarchal societies, females are more risk averse than males (Charness and Gneezy 2012; for evidence suggesting a biologically-informed explanation see, for instance, Brañas-Garza, Galizzi, and Nieboer 2018), this might explain why they tend to use such a strategy to a larger extent than males.

However, while non-*Gitano* females' punishment was strongly modulated by group type—high in the homogeneous and low in the mixed groups—*Gitano* females practically did not punish in either condition. This result may be reflecting a culture-specific differential role of females and males on norm enforcement. Indeed, the finding is consistent with the ethnographic evidence reviewed in Text S1 suggesting that the *Gitano* cultural norms prescribe women to reduce their assertiveness in the presence of (*Gitano*) males, who should ostensibly lead social interactions in such situations. These marked gender roles are far less

prevalent in the majority population. Thus, this result also aligns well with a CGS-related interpretation.

In sum, our results are more consistent with cultural group selection theories and their associated norm-psychology account than with misfiring-based theories for our four research questions. This conclusion does not preclude the importance of punishment as a mechanism for deterrence or that punishers' fitness might be positively affected in some way by relative-standing or reputation gains for example (Raihani and Bshary 2019), but our data indicate that cultural evolution and group selection processes need to be accounted for to explain punishment behavior. However, several findings (in particular, those related to non-*Gitano* punishers in the mixed groups) challenge a radical view of how such processes should translate into behavioral outcomes. These findings in fact raise a number of new questions that deserve further exploration and can help qualify the predictions and interpretation of CGS-related theories.

References

Akerlof, G. A., & Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, *115*(3), 715-753.

Bauer, M., Cahlíková, J., Chytilová, J., & Želinský, T. (2018). Social contagion of ethnic hostility. *Proceedings of the National Academy of Sciences*, *115*(19), 4881-4886.

Benet-Martínez, V., Leu, J., Lee, F., & Morris, M. W. (2002). Negotiating biculturalism cultural frame switching in biculturals with oppositional versus compatible cultural identities. *Journal of Cross-Cultural Psychology*, *33*(5), 492-516.

Bernhard, H., Fischbacher, U., & Fehr, E. (2006). Parochial altruism in humans. *Nature*, *442*(7105), 912-915.

Bowles, S. (2006). Group competition, reproductive leveling, and the evolution of human altruism. *Science*, *314*(5805), 1569-1572.

Boyd, R., & Richerson, P. J. (1988). The evolution of reciprocity in sizable groups. *Journal of Theoretical Biology*, *132*(3), 337-356.

Boyd, R., & Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology*, *13*(3), 171-195.

Boyd, R., Gintis, H., Bowles, S., & Richerson, P. J. (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences*, *100*(6), 3531-3535.

Brañas-Garza, P., Cobo-Reyes, R., & Domínguez, A. (2006). "Si él lo necesita": Gypsy fairness in Vallecas. *Experimental Economics*, *9*(3), 253-264.

Brañas-Garza, P., Espín, A. M., Exadaktylos, F., & Herrmann, B. (2014). Fair and unfair punishers coexist in the Ultimatum Game. *Scientific Reports*, *4*, 6025.

Brañas-Garza, P., Galizzi, M., & Nieboer, J. (2018). Experimental and self-reported measures of risk taking and digit ratio (2D: 4D): evidence from a large, systematic study. *International Economic Review*, *59*(3), 1131-1157.

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup hate?. *Journal of Social Issues*, *55*(3), 429-444.

Brubaker, R., Feischmidt, M., Fox, J., & Grancea, L. (2018). *Nationalist politics and everyday ethnicity in a Transylvanian town*. Princeton University Press.

Charness, G., & Gneezy, U. (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization*, *83*(1), 50-58.

Choi, J. K., & Bowles, S. (2007). The coevolution of parochial altruism and war. *Science*, *318*(5850), 636-640.

Chudek, M., & Henrich, J. (2011). Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Sciences*, *15*(5), 218-226.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. *The Adapted Mind*, 163-228.

Crockett, M. J., Apergis-Schoute, A., Herrmann, B., Lieberman, M. D., Müller, U., Robbins, T. W., & Clark, L. (2013). Serotonin modulates striatal responses to fairness and retaliation in humans. *The Journal of Neuroscience*, *33*(8), 3505-3513.

Curry, O., Whitehouse, H., & Mullins, D. (2019). Is it good to cooperate? Testing the theory of morality-as-cooperation in 60 societies. *Current Anthropology*, *60*(1), 47-69.

de Quervain, D. J., Fischbacher, U., Treyer, V., & Schellhammer, M. (2004). The neural basis of altruistic punishment. *Science*, *305*(5688), 1254.

Delton, A. W., & Krasnow, M. M. (2017). The psychology of deterrence explains why group membership matters for third-party punishment. *Evolution and Human Behavior*, *38*(6), 734-743.

Delton, A. W., Krasnow, M. M., Cosmides, L., & Tooby, J. (2011). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proceedings of the National Academy of Sciences*, *108*(32), 13335-13340.

Dovidio, J. F., Gaertner, S. L., & Saguy, T. (2008). Another view of "we": Majority and minority group perspectives on a common ingroup identity. *European Review of Social Psychology*, *18*(1), 296-330.

Dreber, A., Rand, D. G., Fudenberg, D., & Nowak, M. A. (2008). Winners don't punish. *Nature*, *452*(7185), 348-351.

Espín, A. M., Brañas-Garza, P., Herrmann, B., & Gamella, J. F. (2012). Patient and impatient punishers of free-riders. *Proceedings of the Royal Society of London B: Biological Sciences*, rspb20122043.

Espín, A. M., Exadaktylos, F., Herrmann, B., & Brañas-Garza, P. (2015). Short-and long-run goals in ultimatum bargaining: impatience predicts spite-based behavior. *Frontiers in Behavioral Neuroscience*, *9*, 214.

Exadaktylos, F., Espín, A. M., & Branas-Garza, P. (2013). Experimental subjects are not different. *Scientific Reports*, *3*, 1213

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*(6868), 137-140.

Gächter, S., & Herrmann, B. (2009). Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*(1518), 791-806.

Gächter, S., Renner, E., & Sefton, M. (2008). The long-run benefits of punishment. *Science*, *322*(5907), 1510-1510.

Gamella, J. F., & Álvarez-Roldán, A. (2021). Breaking secular endogamy. A case study of the growth of intermarriage among the Gitanos/Calé of Spain. *Mimeo*.

Gamella, J. F., Martín, E. (2007). Vien avec moi cousine. Le mariage entre cousins germains chez le Gitans Andalous. *Etudes Tsiganes*, 30, 110-151.

Gamella, J. F. (2020). Consanguineous marriages among Andalusian Gitanos/Calé: a genealogical analysis (1925–2006). *Journal of Biosocial Science*, 52(6): 809-831.

Gamella, J. F. (2011). *Historias de éxito. Modelos para reducir el abandono escolar de la adolescencia gitana.* Madrid: IFIIE, Ministerio de Educación.

Gamella, J. F. (2000). *Mujeres gitanas: matrimonio y género en la cultura gitana de Andalucía*. Junta de Andalucía.

Gardner, A., & West, S. A. (2004). Cooperation and punishment, especially in humans. *The American Naturalist*, *164*(6), 753-764.

Gay Blasco, P. (1999). *Gypsies in Madrid: Sex, gender and the performance of identity*. Bloomsbury Academic.

Gil-White, F. J. (2001). Are ethnic groups biological "species" to the human brain? Essentialism in our cognition of some social categories. *Current Anthropology*, *42*(4), 515-553.

Gil-White, F. J. (2004). Ultimatum game with an ethnicity manipulation. In J. Henrich, R. Boyd, S. Bowles, H. Gintis, E. Fehr, and C. Camerer (Eds.) *Foundations of human sociality: Economic experiments and ethnographic evidence from fifteen small-scale societies* (pp. 260-304). Oxford: Oxford University Press

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior*, *24*(3), 153-172.

Goette, L., Huffman, D., & Meier, S. (2006). The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups. *American Economic Review*, *96*(2), 212-216.

Hagen, E. H., & Hammerstein, P. (2006). Game theory and human evolution: A critique of some recent interpretations of experimental games. *Theoretical Population Biology*, *69*(3), 339-348.

Hamilton, W. D. (1964). The genetical evolution of social behaviour. *Journal of Theoretical Biology*, *7*, 1-52.

Hamilton, D. L., Sherman, S. J., & Lickel, B. (1998). Perceiving social groups: The importance of the entitativity continuum. In C. Sedikides, J. Schopler, & C. A. Insko (Eds.), *Intergroup cognition and intergroup behavior* (p. 47–74). Lawrence Erlbaum Associates Publishers.

Henrich, J. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior & Organization*, *53*(1), 3-35.

Henrich, J. (2015). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press.

Henrich, J. (2020). *The weirdest people in the world: How the west became psychologically peculiar and particularly prosperous*. Farrar, Straus and Giroux.

Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., ... & Ziker, J. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science*, *327*(5972), 1480-1484.

Henrich, J., & Henrich, N. (2014). Fairness without punishment: behavioral experiments in the Yasawa Island, Fiji. In *Experimenting with Social Norms: Fairness and Punishment in Cross-Cultural Perspective*, 171-218.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, E., ... & Gwako, E. (2006). Costly punishment across human societies. *Science* 312, 1767–1770.

Henrich, J., & Muthukrishna, M. (2021). The origins and psychology of human cooperation. *Annual Review of Psychology*, *72*, 207-240.

Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science*, *319*(5868), 1362-1367.

Jordan, J. J., McAuliffe, K., & Warneken, F. (2014). Development of in-group favoritism in children's third-party punishment of selfishness. *Proceedings of the National Academy of Sciences*, *111*(35), 12710-12715.

Kardos, P., Leidner, B., Castano, E., & Lickel, B. (2019). The benefits of collective responsibility: How ingroup reputation concern motivates prosociality in intergroup contexts. *European Journal of Social Psychology*, *49*(1), 93-109.

Krasnow, M. M., Cosmides, L., Pedersen, E. J., & Tooby, J. (2012). What are punishment and reputation for?. *PloS ONE*, *7*(9), e45662.

Krupp, D. B., Debruine, L. M., & Barclay P. (2008). A cue of kinship promotes cooperation for the public good. *Evolution & Human Behavior, 29*, 49–55.

Ledyard, J. (1995). Public goods: A survey of experimental research. *The Handbook of Experimental Economics*. Princeton University Press.

Lehmann, L., Rousset, F., Roze, D., & Keller, L. (2007). Strong reciprocity or strong ferocity? A population genetic view of the evolution of altruistic punishment. *The American Naturalist*, *170*(1), 21-36.

Lickel, B., Schmader, T., & Spanovic, M. (2007). Group conscious emotions: The implications of others' wrongdoings for identity and relationships. In J. L. Tracy, R. W. Robins & J. P. Tangney (Eds.), The self-consciouse emotions: Theory and research (pp. 351–370). New York, NY: Guilford Press.

Marlowe, F. W., Berbesque, J. C., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J. C., ... & Tracer, D. (2008). More 'altruistic'punishment in larger societies. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1634), 587-592.

Marques, J., Abrams, D., & Serodio, R. G. (2001). Being better by being right: Subjective group dynamics and derogation of in-group deviants when generic norms are undermined. *Journal of Personality and Social Psychology*, *81*(3), 436.

Marques, J. M., Yzerbyt, V. Y., & Leyens, J. P. (1988). The "black sheep effect": Extremity of judgments towards ingroup members as a function of group identification. *European Journal of Social Psychology*, *18*(1), 1-16.

Martin, E., & Gamella, J. F. (2005). Marriage practices and ethnic differentiation: The case of Spanish Gypsies (1870–2000). *The History of the Family*, *10*(1), 45-63.

Martín, J., Brañas-Garza, P., Espín, A. M., Gamella, J. F., & Herrmann, B. (2019). The appropriate response of Spanish Gitanos: Short-run orientation beyond current socio-economic status. *Evolution and Human Behavior*, *40*(1), 12-22.

Marushiakova, E., & Popov, V. (2007). The Gypsy court in eastern Europe. *Romani Studies*, *17*(1), 67-101.

Mathew, S., & Boyd, R. (2011). Punishment sustains large-scale cooperation in prestate warfare. *Proceedings of the National Academy of Sciences*, *108*(28), 11375-11380.

McDonald, M. M., Navarrete, C. D., & Van Vugt, M. (2012). Evolution and the psychology of intergroup conflict: the male warrior hypothesis. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *367*(1589), 670-679.

McLeish, K. N., & Oxoby, R. J. (2007). Identity, cooperation, and punishment. *IZA Discussion Paper No. 2572.*

McLeish, K. N., & Oxoby, R. J. (2011). Social interactions and the salience of social identity. *Journal of Economic Psychology*, *32*(1), 172-178.

Mendoza, S. A., Lane, S. P. & Amodio, D. M. (2014). For members only: ingroup punishment of fairness norm violations in the ultimatum game. *Social Psychological and Personality Science* 5(6), 662–670.

Mesoudi, A. (2016). Cultural evolution: integrating psychology, evolution and culture. *Current Opinion in Psychology*, *7*, 17-22.

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, *314*(5805), 1560-1563.

Panchanathan, K., & Boyd, R. (2003). A tale of two defectors: the importance of standing for evolution of indirect reciprocity. *Journal of Theoretical Biology*, *224*(1), 115-126.

Pinel, E. C. (1999). Stigma consciousness: the psychological legacy of social stereotypes. *Journal of Personality and Social Psychology*, *76*(1), 114.

Raihani, N. J., & Bshary, R. (2019). Punishment: one tool, many uses. *Evolutionary Human Sciences*, *1*, e12.

Richerson, P. J., & Boyd, R. (2008). *Not by genes alone: How culture transformed human evolution*. University of Chicago Press.

Richerson, P., Baldini, R., Bell, A. V., Demps, K., Frost, K., Hillis, V., ... & Ross, C. (2016). Cultural group selection plays an essential role in explaining human cooperation: A sketch of the evidence. *Behavioral and Brain Sciences*, *39*.

Rusch, H. (2018). Ancestral kinship patterns substantially reduce the negative effect of increasing group size on incentives for public goods provision. *Journal of Economic Psychology, 64, 105-115*.

Rusch, H. (2014). The evolutionary interplay of intergroup conflict and altruism in humans: a review of parochial altruism theory and prospects for its extension. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1794), 20141539.

Sääksvuori, L., Mappes, T., & Puurtinen, M. (2011). Costly punishment prevails in intergroup conflict. *Proceedings of the Royal Society of London B: Biological Sciences*, rspb20110252.

San Román, T. (2010). *La diferencia inquietante: viejas y nuevas estrategias culturales de los gitanos*. Siglo XXI de España Editores.

Schiller, B., Baumgartner, T., & Knoch, D. (2014). Intergroup bias in third-party punishment stems from both ingroup favoritism and outgroup discrimination. *Evolution and Human Behavior*, *35*(3), 169-175.

Schulz, J. F., Bahrami-Rad, D., Beauchamp, J. P., & Henrich, J. (2019). The Church, intensive kinship, and global psychological variation. *Science*, *366*(6466), eaau5141.

Sell, A., Tooby, J., & Cosmides, L. (2009). Formidability and the logic of human anger. *Proceedings of the National Academy of Sciences*, *106*(35), 15073-15078.

Shelton, J. N. (2003). Interpersonal concerns in social encounters between majority and minority group members. *Group Processes & Intergroup Relations*, *6*(2), 171-185.

Shinada, M., Yamagishi, T., & Ohmura, Y. (2004). False friends are worse than bitter enemies: "Altruistic" punishment of in-group members. *Evolution and Human Behavior*, *25*(6), 379-393.

Sigmund, K. (2007). Punish or perish? Retaliation and collaboration among humans. *Trends in Ecology & Evolution*, *22*(11), 593-600.

Soltis, J., Boyd, R., & Richerson, P. J. (1995). Can group-functional behaviors evolve by cultural group selection?: An empirical test. *Current Anthropology*, *36*(3), 473-494.

Tabibnia, G., & Lieberman, M. D. (2007). Fairness and cooperation are rewarding. *Annals of the New York Academy of Sciences*, *1118*(1), 90-101.

Tajfel, H. (1974). Social identity and intergroup behaviour. *Social Science Information/sur les Sciences Socials*, *13*(2), 65-93.

Tooby, J., & Cosmides, L. (1990). The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology and Sociobiology*, *11*(4), 375-424.

Tropp, L. R., & Pettigrew, T. F. (2005). Relationships between intergroup contact and prejudice among minority and majority status groups. *Psychological Science*, *16*(12), 951-957.

Walker, R. S. (2014). Amazonian horticulturalists live in larger, more related groups than hunter–gatherers. *Evolution and Human Behavior*, *35*(5), 384-388.

Weyrauch, W. O. (Ed.) (2001). *Gypsy law: Romani legal traditions and culture*. University of California Press.

Yudkin, D. A., Van Bavel, J. J., & Rhodes, M. (2019). Young children police group members at personal cost. *Journal of Experimental Psychology: General*. 149 (1), 182-191.

Authors' contributions

AME designed and conducted the experiment, performed the statistical analyses, and wrote the paper; PBG designed and conducted the experiment and wrote the paper; BH designed the experiment and wrote the paper; JFG designed and conducted the experiment and wrote the paper; JM performed the statistical analyses and wrote the paper.

Competing interests

The authors declare no conflict of interest.

Funding

Financial support from the Spanish Ministry of Science and Innovation (PGC2018-093506-B-I00, ECO2013-44879-R) and the Regional Government of Andalusia (PY18-FR-0007, P11-SEJ-8286 and P12-SEJ-1436) is gratefully acknowledged. Antonio Espín acknowledges funding from the European Union's Horizon 2020 research and innovation program under the

Marie Skłodowska-Curie grant agreement No. 754446 and UGR Research and Knowledge Transfer Fund – Athenea3i.

Acknowledgements

We thank F. Exadatyklos, J. Henrich, R. Hernán-González, H. Kaplan, E. Molis, J. C. Perales, P. Román, H. Rusch, and J. Schulz for their helpful comments; J. A. Abril, A. Cortés, J. Martín, E. M. Muñoz, L. A. Palacios, M. Parravano, L. E. Pedauga, A. Quesada, M. Román, and J. F. Ruano for their excellent research assistance; and the audience of seminars/talks at Durham, East Anglia, Exeter, Granada, Lancaster, Loyola Andalucía, Middlesex, Navarra, SEJyD-2016, and Virginia. Special thanks are due to the authorities and inhabitants of Benalúa, Darro, Deifontes, Iznalloz, and Pedro Martínez.

Main Text Figures

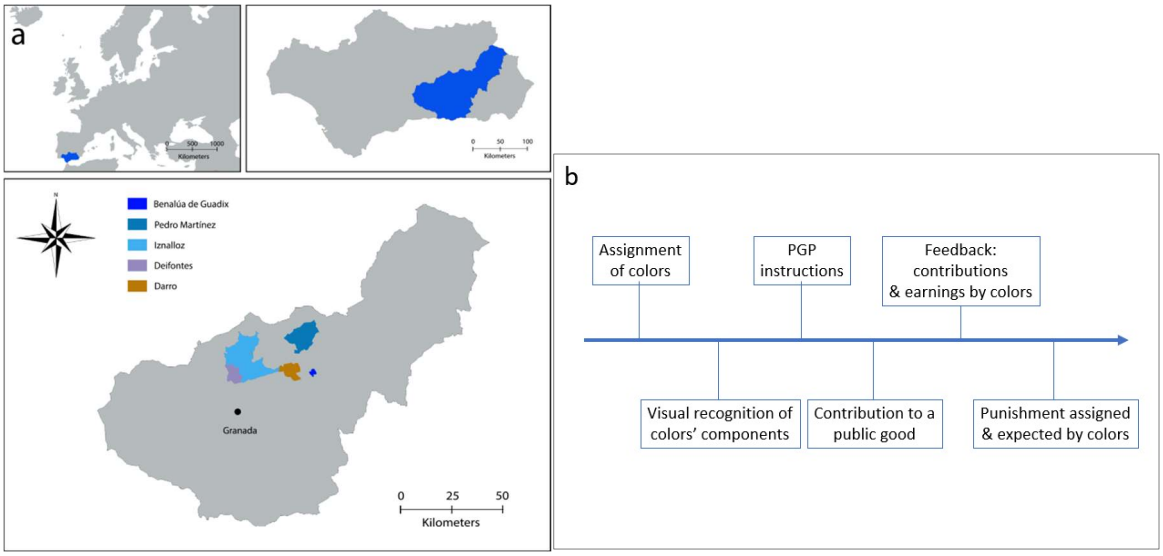

Figure 1. *Panel (a) Five semi-rural towns in southern Spain (Granada, Andalusia): Benalúa de Guadix, Darro, Deifontes, Iznalloz, and Pedro Martínez. Panel (b) Structure of the experiment.*

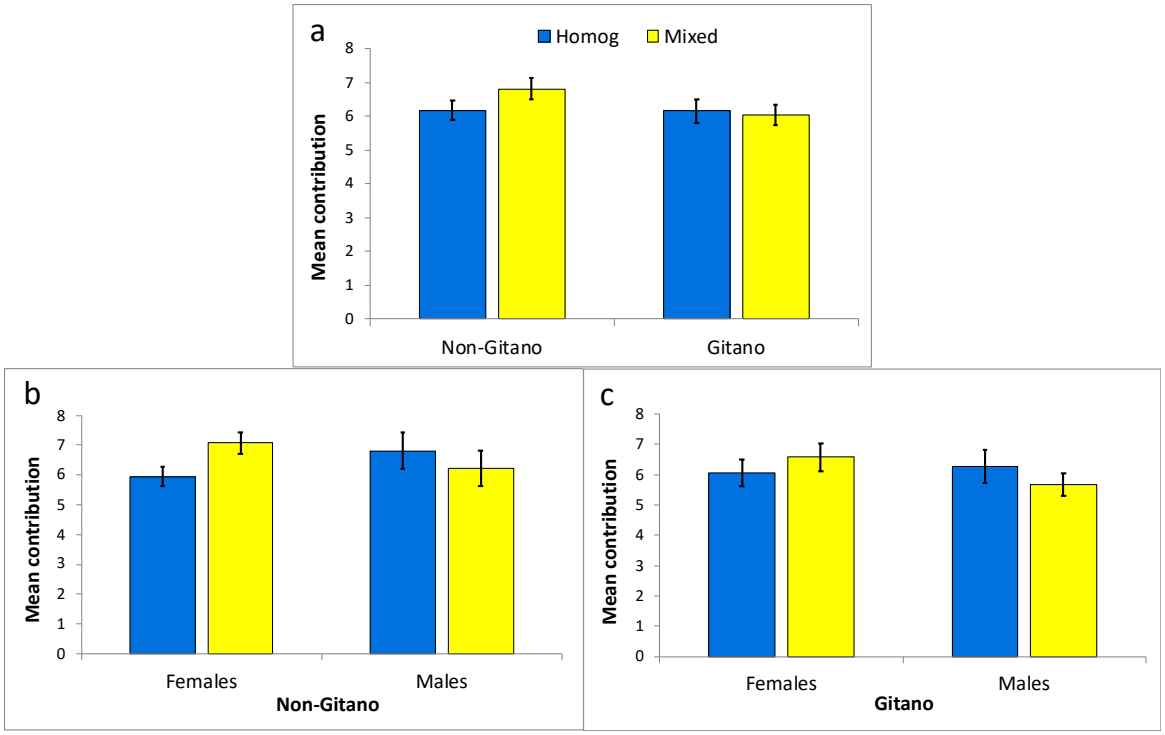

Figure 2. *Mean contributions in homogeneous and mixed conditions. Panel (a) displays the data broken down by ethnicity. Panels (b) and (c) display the data broken down by gender for non-Gitanos and Gitanos, respectively. Error bars represent standard error of the mean.*

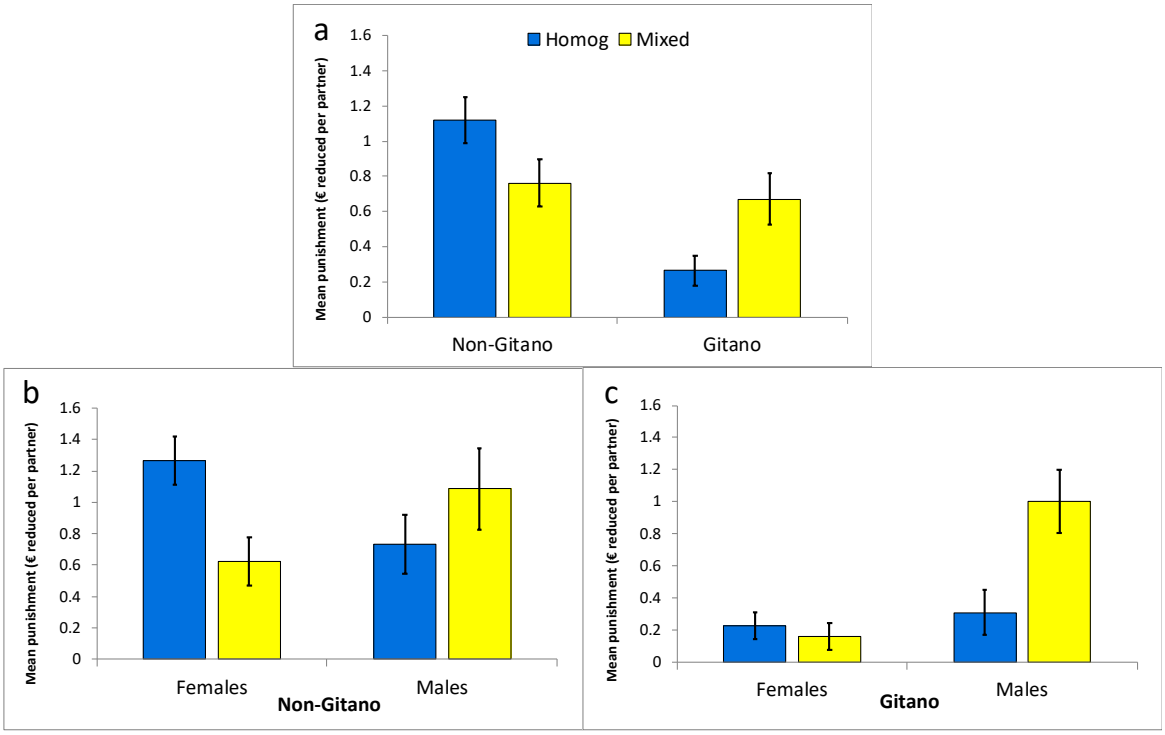

Figure 3. *Mean aggregate punishment in homogeneous and mixed conditions. Panel (a) displays the data broken down by ethnicity. Panels (b) (non-Gitanos) and (c) (Gitanos) display the data broken down by ethnicity and gender. Error bars represent robust standard error of the mean clustered at the group level.*

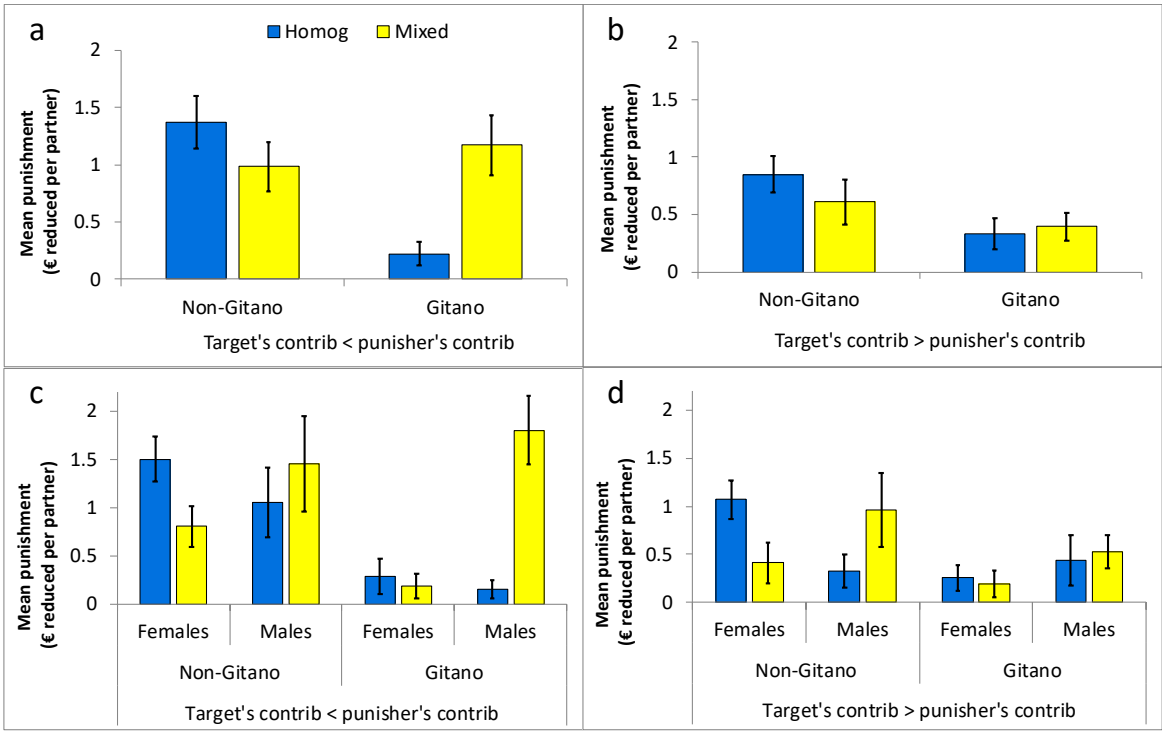

Figure 4. *Mean altruistic and antisocial punishment in homogeneous and mixed conditions. Panels (a) (altruistic punishment) and (b) (antisocial punishment) display the data broken down by punisher's ethnicity. Panels (c) (altruistic punishment) and (d) (antisocial punishment) display the data broken down by punisher's ethnicity and gender. Error bars represent robust standard error of the mean clustered at the group level.*

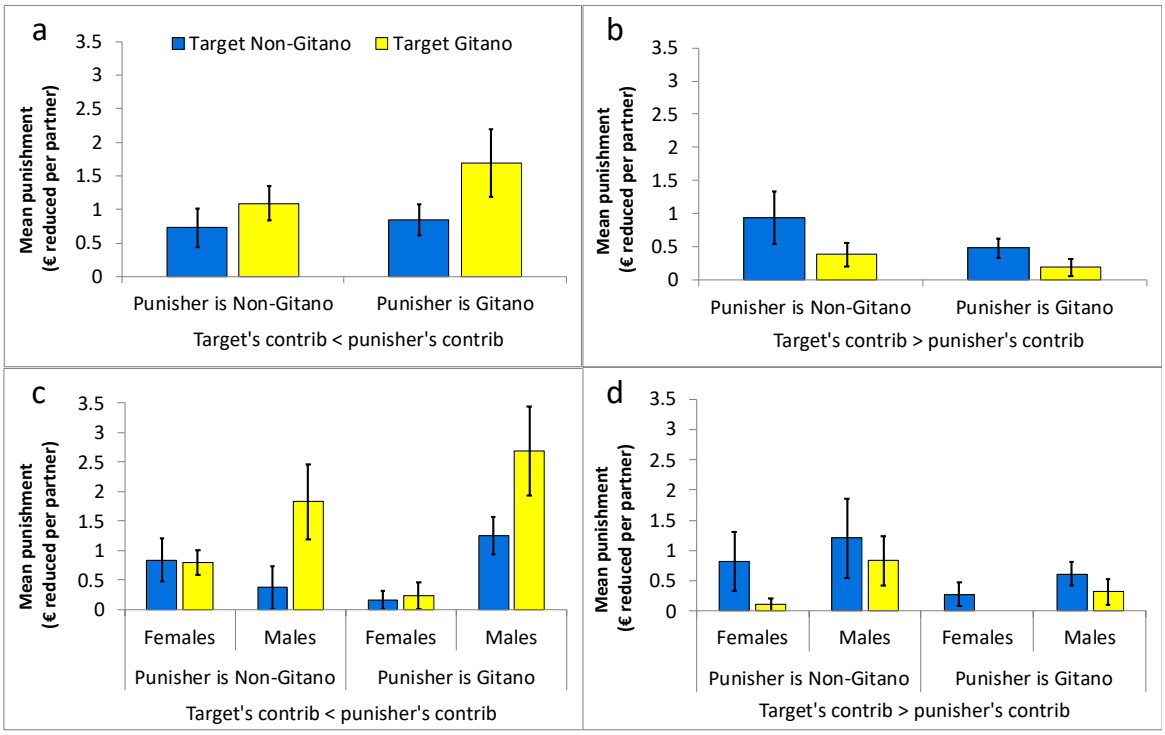

Figure 5. *Mean punishment on Gitano and non-Gitano targets in mixed groups. Panels (a) (altruistic punishment) and (b) (antisocial punishment) display the data broken down by punisher's ethnicity. Panels (c) (altruistic punishment) and (d) (antisocial punishment) display the data broken down by punisher's ethnicity and gender. Error bars represent robust standard error of the mean clustered at the group level.*

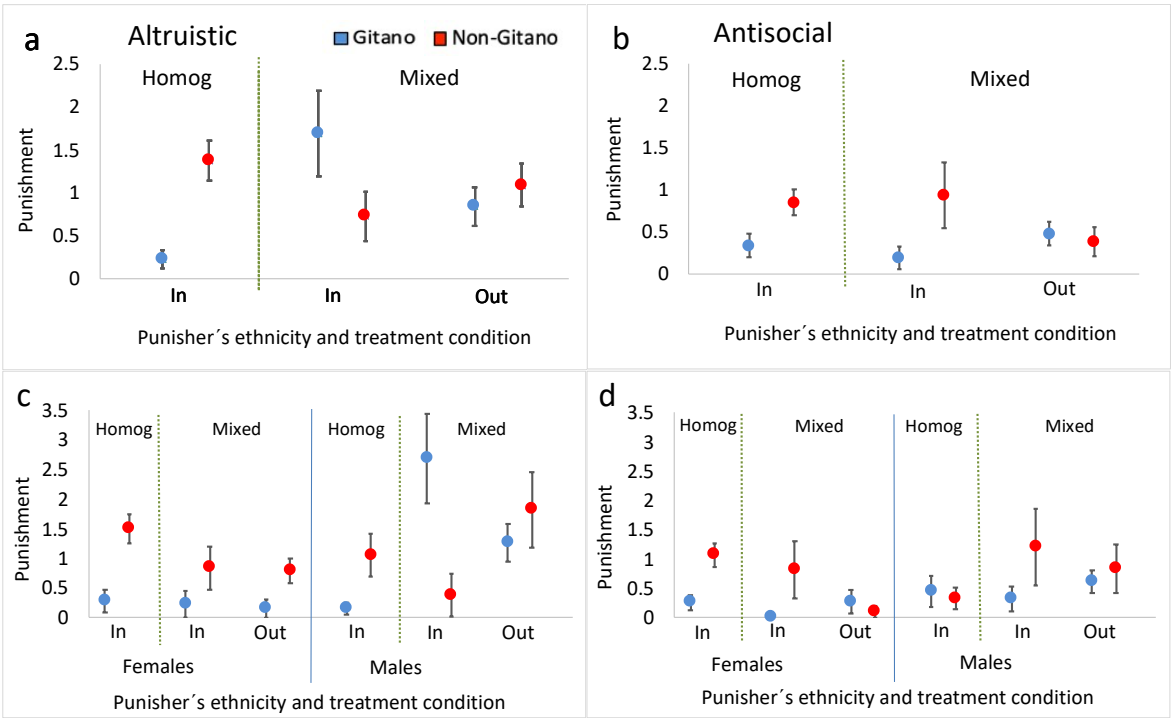

Figure 6. *Mean altruistic and antisocial punishment targeted at ingroups and outgroups. Panels (a) (altruistic punishment) and (b) (antisocial punishment) display the data broken down by punisher's ethnicity and treatment condition (ingroup-homogeneous, ingroup-mixed and outgroup-mixed). Panels (c) (altruistic punishment) and (d) (antisocial punishment) display the data broken down by punisher's ethnicity, treatment condition, and punisher's gender. Error bars represent robust standard error of the mean clustered at the group level.*

Supplementary Materials

for

Group-level ethnic composition influences altruistic punishment: public goods experiments among *Gitanos* and non-*Gitanos* in southern Spain

Text S1. A short overview on Spanish *Gitanos* and norm-psychology hypotheses

The *Gitanos* or *Calé*[5] are an ethnocultural minority that lives today in all Spanish regions. They are related to other Romani groups in Europe and America with whom they seem to share a remote origin from an "initial founder population" that moved westwards from the Indian subcontinent over one thousand years ago (Mendizabal et al. 2012). All these groups, however, have adapted to the surrounding groups with whom they have lived and today show some traits of familial resemblance and considerable cultural heterogeneity (Matras 2015; Piasere 2004; Fraser 1992). Even those who preserve articulated dialects of Romani language (Matras 2002) are bilingual, and thus bicultural. The *Gitanos* come from the first Romani migrations into Western Europe, which ended in the second half of the 15th century (Pym 2007; Leblon 1985). Their lifeways are the product of a long coexistence and exchange with local Spanish populations. Life in common has been marked by persecution, segregation, and discrimination, but also by cooperation and hybridization (Pym 2007; Gómez Alfaro 1998; 1999; Leblon 1985; Gamella 2011; Gamella et al. 2014b).

[5] Most Spanish Romani people call themselves *Gitanos* in both private and public settings. Minority leaders also use the term to name public institutions, such as the *Instituto de Cultura Gitana*. The first Romani groups reaching Spain in the fifteenth century were called "*Egyptanos*", as they were believed to have originated in Egypt. *Gitano* is thus synonymous with the English term "Gypsy." Many Romani activists and intellectuals reject this exonym as derogatory and prefer to be identified by their own denominations, such as Roma, Sinti, Kalé, etc. Some leaders of the growing international Roma movement and some EU authorities defend the term "Roma" for all Romani groups. We rarely heard the term "Roma" in our encounters with *Gitano* people. In Spain, *Gitanos* also refer to themselves as *Calé* (plural of *Caló*, black in Romani), but less frequently.

In this sense, *Gitanos* of Spain are often portrayed as an example of successful integration. Arguably, their treatment and living conditions are relatively favorable compared to large Romani populations living in other European societies, particularly those of Central and Eastern Europe. (For instance, George Soros, the business magnate and Roma advocate and philanthropist "called upon Spain to lead Europe in bettering the conditions of the Roma" [Peiró 2012:ix]. Similar claims have been expressed often in the international mass media.) But the rosy view of the lot of the Spanish Romani is often exaggerated and downplays the discrimination and disadvantage many *Gitano* men and women still suffer in labor, income, education, and even daily life encounters (Álvarez-Roldán et al. 2018). It is true, however, that since 1977, when the new political context brought about democracy and decentralization of the Spanish state, there have been clear improvements in their access to health care, education, and housing, but not without conflicts and rejection by local majorities.

Today, most *Gitanos* are proud of their ethnic identity, although they consider themselves autochthonous Spaniards, especially in face of the large number of foreign economic immigrants who moved to Spain in the last two decades and increased the country's ethnic and cultural diversity. *Gitanos* speak the languages and dialects of the regions where they live and have lost most of their old trades and occupations. They have, however, developed other differences in religious expression and mobilization or in gender and marriage rituals, as well as in reproductive patterns, to construct and vindicate their shared identity (Gay Blasco 1999; Cantón 2010, 2020; Gamella et al. 2013, 2014a, 2014b). *Gitanos*' identity often shows elements of an "oppositional identity" built in opposition or in contrast to the dominant majority culture and associated with the status of involuntary minority (Ogbu and Simons 1998). But *Gitanos* have contributed much to Spanish culture and folklore. Perhaps in no other part of Europe has such a cultural fusion occurred as in Spain, especially in Andalusia, where many of the symbols and practices that identify the region to the world (such as

flamenco singing and dancing) have a crucial *Gitano* component (Leblon 2003; Pasqualino 1998).

Almost all Spanish *Gitanos* are sedentary; they have lived in the same towns and counties for generations and often have a strong attachment to their places of birth or residence, defining themselves as Andalusians, Catalans, or even *Sevillanos* and *Granadinos*. Informed estimates of the size of the *Gitano* population put it in the range of 500,000 to 600,000, around 1.5% of the total Spanish population (FSG 2008). Although in some locations, mainly in the southern region of Andalusia where about 40% of the Spanish *Gitanos* live (even though Andalusia has less than 20% of the total Spanish population), *Gitanos* represent a particularly high fraction of the population. We conducted our study in an area of eastern Andalusia. This geographical area was chosen due to its high concentration of *Gitanos*, thus allowing the recruitment of a sufficient number of members of this ethnicity for our study. In the five towns hosting the experiments, *Gitanos* account for about 25.6% of the population on average (range: 20.0%–41.4%), that is, about 3,970 over a total of 15,490 inhabitants according to our estimates for 2007.

Some *Gitano* cultural traits are essential for understanding their social behavior and peer punishment in particular. Such traits are mainly associated with social organization and gender roles. We summarize their differential characteristics in the following lines and develop hypotheses about how some of these cultural traits might translate, as proximate-level explanations, into observed behavior in the experiment.

Social organization and "the family"

Even considering the growing heterogeneity of *Gitanos*, their social universe is largely based on kinship and marriage relations. Their main social networks are family networks, which

tend to be larger, denser, and more complex and multifunctional than those of their non-*Gitano* neighbors (or *Payos*, as *Gitanos* often refer to them).

For *Gitanos* today, their most important institution is "the family." The particular notion of family among the *Gitano* population encompasses many different meanings, which can be summarized across two levels. First, compared to non-*Gitanos*, *Gitanos* display relatively smaller stress in the household or co-resident domestic unit and a more general understanding of the "closest family" as including a network of households formed by close kinship links. Considering the different moments in the developmental cycle of domestic units, it is possible to find, for instance, that a specific couple and their children gravitate heavily and almost daily towards the husband's parents. Thus, a patri-virilocal bias strengthens the patrilineal ideology sustained primarily by males (Gay Blasco 1999; Martín and Gamella 2005; Gamella and Martín 2007; Gamella 2011). Second, kin networks include a larger number of people due to several processes that differ from the majority at large: in particular, (i) higher fertility leading to a larger number of siblings and, in turn, aunts-uncles, cousins, second cousins, etc.; and (ii) higher consanguinity in marriage that generates a multiplicity of links between members of any network, as well as higher network homogeneity, although in the last decades the heterogeneity of *Gitano* families may be increasing (Gamella and Álvarez-Roldán 2021).

Inbreeding has indeed been strikingly common among *Gitanos*, who show a marked preference to marry "known," compatible, and "good" people from reliable interrelated kin networks. This does not stem only from geographic isolation or inheritance rules and patrimonial strategies. Rather, it is more the result of social isolation or segregation, as well as a marked cultural preference for endogamy (Gamella 2020).

It has long been argued that in premodern or "traditional" societies kinship "provides […] an organising medium of trust relations." As such, "kinspeople can usually be relied upon to

meet a range of obligations more or less regardless of whether they feel personally sympathetic towards the specific individuals involved", while in modern societies relationships of trust have been replaced by "friendship or sexual intimacy as a means of stabilising social ties" (Giddens 1990:101–102). The dominant idea is that modernity implies isolation from kin networks and individuals confront each other as separate entities "divorced from their kinship and family units" (Finkler et al. 2001:236). This varies across countries, however (Schulz et al. 2019). Precisely, Spain as well as other southern European countries are usually portrayed as "familial" societies, where family bonds and support are relatively prominent, and individualism is somehow limited by family obligations (Reher 1998). Therefore, the distinction between *Gitanos* and Spaniards at large in this regard might be considered as a question of degree rather than as an absolute one. But the density and intensity of kin bonds often generate a differential institutional setup and affect the interpretative lens shared by local *Gitanos*.

Inbreeding is much more common among *Gitanos* than among Spaniards at large and has shown both a distinctive character and evolution. Although Spain once had some of the highest levels in Europe, inbreeding began to fall in the 1950s and, in following decades, the fall was so rapid that consanguineous marriages have become as rare as in other Western countries (Fuster and Colantonio 2002, 2004; Calderón et al. 2009). Within *Gitano* communities, however, inbreeding has been and remains widespread. According to recent estimates based on genealogical reconstruction for the period 1925–2006 (Gamella 2020), in 22 contiguous localities in the area where this study was conducted more than half (54.8%) of all *Gitano* marriages are among relatives, with close-kin consanguineous marriages (up to second cousins) averaging 28.7%. An estimation that can be compared to the measures reported in studies using interviews or other synchronic research methods yields average inbreeding coefficients (Wright's F) of about 11.3 ($x10^{-3}$), levels never found in Spain and

much less so recently. This value is rather conservative, however, and may underestimate the actual F by more than 30% in this population. These are among the highest rates of inbreeding found in any European population, including the most inbred of Spanish isolates (Gamella 2020). In the same area, aggregate consanguinity rates for the overall population (including *Gitanos* and non-*Gitanos*) reached a maximum of around 7.4% between 1920 and 1936, with corresponding F coefficients ranging from 2.4 to 2.7 ($x10^{-3}$). Since the 1960s, the rates of consanguinity and inbreeding have decreased rapidly (Gamella and Núñez-Negrillo 2019). Note that recent comparable estimates for small-scale societies of hunter-gatherer and horticulturalists report average F values well below 2 ($x10^{-3}$) and 10 ($x10^{-3}$), respectively (Walker 2014; Walker and Bailey 2014). Given the strong correlation between coefficients of inbreeding and mean relatedness (Hamilton's r) of groups (Walker 2014), these data demonstrate that Romani people of this area are highly genetically related on average, even compared with people from small-scale societies. Multiple consanguinity is the norm among *Gitanos*: couples are linked by several bonds and share many ancestors; a product of a pattern of inbreeding sustained over many generations. However, these patterns are changing and the rate of intermarriage between *Gitanos* and non-*Gitanos* is increasing, particularly in some local communities (Gamella and Álvarez-Roldán 2021).

In sum, even in a region where consanguineous marriages had been important, inbreeding among *Gitanos* shows a particularly high intensity and permanence, as it is the product of a strong cultural preference and not only of geographical isolation and poverty. Hence, it is somehow reasonable that *Gitanos* spread that sense of kin to the whole community: "here we all are family"; "all *Gitanos* are related, they share some blood, at least a drop of blood for sure"; "distant but relatives". Neighbors, friends, and partners are often family as well.

The enforcement of norms—a norm-psychology hypothesis

Regarding norm-enforcement institutions, some Romani groups have formal conflict resolution processes and tribunals. *Gitano* people, however, use more informal systems of justice and adjudication of rights to avoid the escalation of violence and blood feuds (San Román 1986, 2010). Respected elders, typically men (*hombres de razón* or *hombres de respeto*: "men of reason" or "men of respect"), are often asked to mediate. Affinal kin relationships may also limit the extent and seriousness of conflicts, which have been recurrent and feared. Still today a serious conflict (a death) may imply the abandonment of their residences by several hundred of the closest kin of the accused.

Notwithstanding, both male and female *Gitanos*, but in different socio-political spheres, display a comparatively strong sense of individual autonomy (Gamella 2000, 2011) which, added to the possibility of escalation of conflict between families, may restrict the role of decentralized overt sanctioning unless key norms are transgressed (Piasere 2012; Matras 2015; San Román 2010; Gay Blasco 1999; Álvarez-Roldán et al. 2018). This culture of liberty or resistance, possibly related to the avoidance of conflict between *Gitano* families, should be associated with a low willingness to punish in homogeneous groups if cultural differences are translated into game play as predicted by a norm-psychology account. An earlier study with a sample of Spanish *Gitanos* provides preliminary support for this prediction. Brañas-Garza et al. (2006) used ultimatum game experiments to examine sharing and punishment behavior in anonymous one-shot bilateral interactions between *Gitanos* in Vallecas, Madrid. Most of them did not express any willingness to punish stingy co-ethnics (but see Espín et al. 2012, 2015 for combined evidence suggesting that the psychology underlying the rejection of low offers in the ultimatum game may differ from that underlying altruistic punishment in the PGP). Furthermore, a common rationale of *Gitanos* who were unwilling to reject unfair, even zero, offers was, "What if (s)he needs the money?". This suggests that sporadic acts of

uncooperativeness carried out by *Gitanos* may not per se be considered by other *Gitanos* as deserving peer punishment; solidarity and forgiveness might be the intuitive response.

Gender roles—a norm-psychology hypothesis

In general, *Gitanos* are portrayed as a group that sustains relatively conservative or patriarchal gender relationships, where women are subordinated to fathers and brothers when they are single, and to their husbands and husband's family when married (San Román 2010; Gay Blasco 1999). Care of children, family members, and the sick are generally seen as women's primordial tasks, but in this regard there is only a degree of difference with non-*Gitanos* of this area.

However, the considerable agency developed by *Gitano* women in their daily lives, both in the domestic and public realms, is rarely considered. It is often *Gitano* women who confront authorities in administrative matters and in the defense of their rights to housing, education or public benefits. But they do that somehow as in delegation by their husbands and partners; it is part of their accepted gender roles. In confrontational encounters judged as impersonal, *Gitano* women can be very assertive and their attitudes are often seen as inadequate by majority standards, as if they were not following the same patterns of modesty and good manners of middle-class Spaniards (Gamella 2000, 2011). This supposed lack of accommodation to their subordinate status is part of the generalized anti-*Gitano* bias that reflects important majority norms; a process also found with respect to anti-Roma bias in Eastern Europe (Kende et al. 2017).

But in personal interactions, or in front of *Gitano* people, the presence of males in public encounters somehow transforms the ways most *Gitano* women will voice their concerns and pursue their interests. There exists a number of principles that *Gitano* women must typically follow in these cases: e.g., "never let him lose face in public" or "never contradict him or the

elders publicly". If women decide or influence family decisions, as they often do, their role has to be more private than public, more by applying reason than violence (Gamella 2000; Gamella and Martín 2007). In this sense, while gossiping is a fundamental weapon in the hands of women, violence is seen as the prerogative of males in extreme circumstances (Gay Blasco 1999; San Román 2010). There is obviously much variation among individuals and couples in these gender arrangements and age may also play an important moderating role, but this norm clearly differs with respect to the majority population. Following the norm-psychology account, this cultural difference is hypothesized to be reflected in game behavior in that *Gitano* females should be more reluctant (than non-*Gitano* ones and males in general) to punish others in either condition of the experiment given that *Gitano* males are always present.

References

Álvarez-Roldán, A., Parra, I., & Gamella, J. F. (2018). Reasons for the underachievement and school drop out of Spanish Romani adolescents. A mixed methods participatory study. *International Journal of Intercultural Relations*, *63*, 113-127.

Brañas-Garza, P., Cobo-Reyes, R., & Domínguez, A. (2006). "Si él lo necesita": Gypsy fairness in Vallecas. *Experimental Economics*, *9*(3), 253-264.

Calderón, R., Aresti, U., Ambrosio, B., & González-Martín, A. (2009). Inbreeding Coefficients for X-linked and Autosomal Genes in Consanguineous Marriages in Spanish Populations: The Case of Guipúzcoa (Basque Country). *Annals of Human Genetics*, *73*(2), 184-195.

Cantón, M. (2010) “Gypsy Pentecostalism, ethnopolitical uses and construction of belonging in the south of Spain”. *Social Compass (International Review of Sociology of Religion)*, Vol. 57, nº 2, pp. 253-267.

Cantón, M. (2020). *Evangelical Gypsies in Spain:" The Bible is our Promised Land"*. Lexington: Lexington Books.

Espín, A. M., Brañas-Garza, P., Herrmann, B., & Gamella, J. F. (2012). Patient and impatient punishers of free-riders. *Proceedings of the Royal Society of London B: Biological Sciences*, rspb20122043.

Espín, A. M., Exadaktylos, F., Herrmann, B., & Brañas-Garza, P. (2015). Short-and long-run goals in ultimatum bargaining: impatience predicts spite-based behavior. *Frontiers in Behavioral Neuroscience*, *9*, 214.

Finkler, K., Dolgin, J., Franklin, S., Gusterson, H., Hadler, N. M., Evans, J. P., ... & Finkler, K. (2001). The kin in the gene: the medicalization of family and kinship in American society. *Current Anthropology*, *42*(2), 235-263.

Fraser, A. (1992). *The Gypsies: The people of Europe*. Oxford: Blackwell Publishers.

FSG (2008) *Mapa sobre vivienda y comunidad gitana en España, 2007*. Madrid: Fundación Secretariado Gitano.

Fuster, V., & Colantonio, S. (2004). Socioeconomic, demographic, and geographic variables affecting the diverse degrees of consanguineous marriages in Spain. *Human Biology*, *76*(1), 1-14.

Fuster, V., & Colantonio, S. E. (2002). Consanguinity in Spain: socioeconomic, demographic, and geographic influences. *Human Biology*, 301-315.

Gamella, J. F. (2000). *Mujeres gitanas: matrimonio y género en la cultura gitana de Andalucía*. Junta de Andalucía.

Gamella, J. F. (2011). *Historias de éxito. Modelos para reducir el abandono escolar de la adolescencia gitana.* Madrid: IFIIE, Ministerio de Educación.

Gamella, J. F., & Álvarez-Roldán, A. (2021). Breaking secular endogamy. A case study of the growth of intermarriage among the Gitanos/Calé of Spain. *Mimeo*.

Gamella, J. F., Carrasco-Muñoz, E. M., & Núñez Negrillo, A. M. (2013). Oculocutaneous albinism and consanguineous marriage among Spanish *Gitanos* or Calé–A study of 83 cases. *Collegium Antropologicum*, *37*(3), 723-734.

Gamella, J. F., Carrasco-Muñoz, E. M., & Quesada Garrido, A. (2014a). New methodological approaches in the anthropological demography of Romani groups. An example from the study of the evolution of the infant and child mortality of the *Gitanos* or Calé of Spain (1871-2007). *Sociologia: Revista da Faculdade de Letras da Universidade Do Porto*, *4*, 175-204.

Gamella, J. F., Gómez Alfaro, A., & Pérez Pérez, J. (2014b). Nominal Assimilation: The Ethnic and National Identities of the *Gitanos* or Calé of Spain as Shown by their Surnames in the 1783–1785 Census. *Names*, *62*(3), 147-164.

Gamella, J. F., Martín, E. (2007). Vien avec moi cousine. Le mariage entre cousins germains chez le Gitans Andalous. *Etudes Tsiganes*, 30, 110-151.

Gamella, J.F. & Núñez-Negrillo, A.M. (2019). The Evolution of Consanguineous Marriages in the Archdiocese of Granada, Spain (1900-1979). *Human Biology*, *90*(2): 97-114.

Gamella, J. F. (2020). Consanguineous marriages among Andalusian Gitanos/Calé: a genealogical analysis (1925–2006). *Journal of Biosocial Science*, 52(6): 809-831.

Gay Blasco, P. (1999). *Gypsies in Madrid: Sex, gender and the performance of identity*. Bloomsbury Academic.

Giddens, A. (1990). *The consequences of modernity*. Cambridge: Polity.

Gómez Alfaro A. (1999). Tipologías. Matrimonios mixtos y mestizajes gitanos en los censos históricos andaluces. *Demófilo*, *30*, 30–52.

Gómez Alfaro, A. (1998). Gitanos de Priego de Córdoba: entre el acoso y la supervivencia. *Legajos. Cuadernos de Investigación Histórica del Sur de Córdoba*, *1*, 61-72.

Kende, A., Hadarics, M., & Lášticová, B. (2017) Anti-Roma attitudes as expressions of dominant social norms in Eastern Europe. *International Journal of Intercultural Relations*, 60: 12-27. DOI: doi.org/10.1016/j.ijintrel.2017.06.002

Leblon, B. (1985). *Les Gitans d'Espagne*. Paris: Presses Universitaires de France.

Leblon, B. (2003). *Gypsies and flamenco: the emergence of the art of flamenco in Andalusia* (Vol. 6). Univ of Hertfordshire Press.

Martin, E., & Gamella, J. F. (2005). Marriage practices and ethnic differentiation: The case of Spanish Gypsies (1870–2000). *The History of the Family*, *10*(1), 45-63.

Matras, Y. (2002). *Romani: A linguistic introduction*. Cambridge University Press.

Matras, Y. (2015). *The Romani Gypsies*. Harvard University Press.

Mendizabal, I., Lao, O., Marigorta, U. M., Wollstein, A., Gusmão, L., Ferak, V., ... & Kučinskas, V. (2012). Reconstructing the population history of European Romani from genome-wide data. *Current Biology*, *22*(24), 2342-2349.

Ogbu, J. U., & Simons, H. D. (1998). Voluntary and involuntary minorities: a cultural-ecological theory of school performance with some implications for education. *Anthropology & Education Quarterly*, *29*(2), 155-188.

Pasqualino, C. (1998). *Dire le chant. Les Gitanes flamencos d'Andalousie*. Paris: CNRS Editions.

Peiró, E. W. (2012). *White Gypsies: Race and stardom in Spanish musical films*. University of Minnesota Press.

Piasere, L. (2004). *I rom d'Europa: una storia moderna*. GLF editori Laterza.

Piasere, L. (2012). *Scenari dell'antiziganismo.* Firenze: SEID Editori.

Pym, R. (2007). *The Gypsies of early modern Spain, 1425-1783*. London: Palgrave Macmillan.

Reher, D. S. (1998). Family ties in Western Europe: persistent contrasts. *Population and Development Review*, *24*(2), 203-234.

San Román, T. (1986). *Entre la marginación y el racismo. Reflexiones sobre la vida de los gitanos*. Madrid: Alianza Universidad.

San Román, T. (2010). *La diferencia inquietante: viejas y nuevas estrategias culturales de los gitanos*. Siglo XXI de España Editores.

Schulz, J. F., Bahrami-Rad, D., Beauchamp, J. P., & Henrich, J. (2019). The Church, intensive kinship, and global psychological variation. *Science*, *366*(6466), eaau5141.

Walker, R. S. (2014). Amazonian horticulturalists live in larger, more related groups than hunter–gatherers. *Evolution and Human Behavior*, *35*(5), 384-388.

Walker, R. S., & Bailey, D. H. (2014). Marrying kin in small-scale societies. *American Journal of Human Biology*, *26*(3), 384-388.

Supplementary Figures and Tables

Figure S1. *Contribution decision card (Yellow #1 participant example; translated from Spanish)*

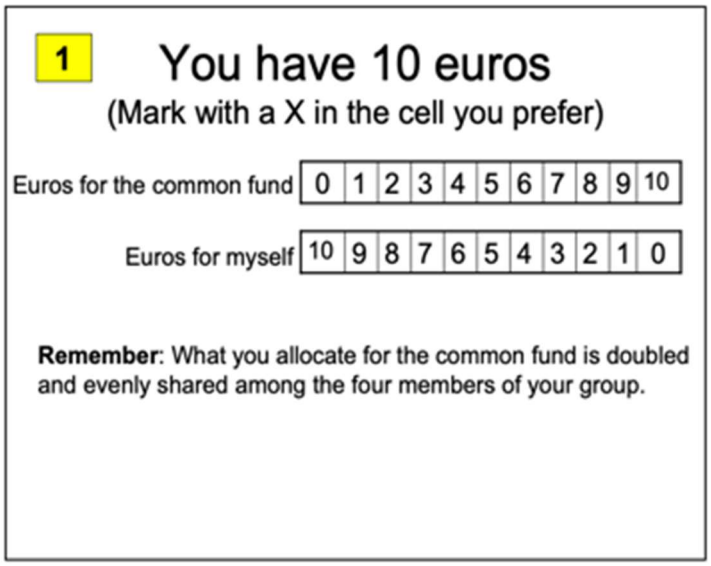

Figure S2. *Punishment decision card (Yellow #1 participant example; translated from Spanish)*

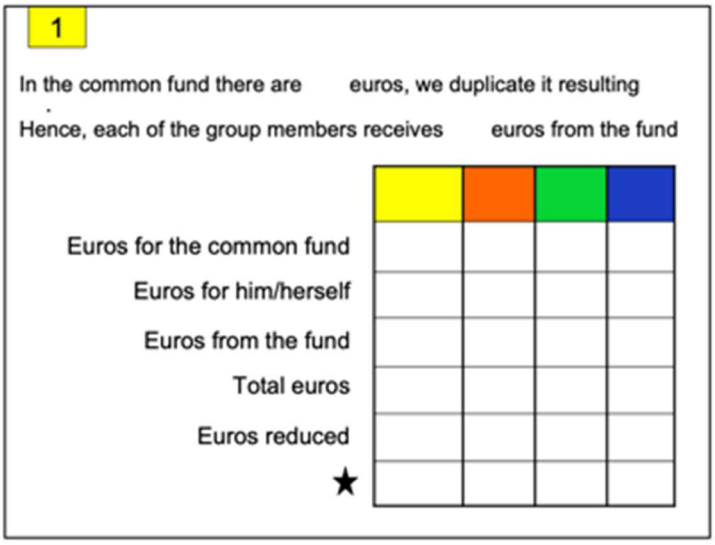

Table S1. *The determinants of contributions to the public good*

depvar: *contribution*	(1a)	(2a)	(3a)	(4a)	(5a)	(1b)	(2b)	(3b)	(4b)	(5b)
gitano	-0.361	0.003	-0.145	-0.313	0.116	-0.270	0.065	0.011	-0.248	0.217
	(0.321)	(0.456)	(0.399)	(0.319)	(0.549)	(0.417)	(0.524)	(0.488)	(0.414)	(0.601)
mixed	0.306	0.639	0.321	0.921**	1.134**	0.388	0.694	0.417	0.971**	1.174**
	(0.312)	(0.423)	(0.313)	(0.384)	(0.483)	(0.323)	(0.428)	(0.324)	(0.387)	(0.482)
male	-0.181	-0.156	0.083	0.605	0.867	-0.131	-0.110	0.225	0.626	0.976
	(0.333)	(0.336)	(0.490)	(0.486)	(0.686)	(0.342)	(0.344)	(0.500)	(0.495)	(0.709)
gitano X mixed		-0.733			-0.611		-0.682			-0.574
		(0.620)			(0.799)		(0.622)			(0.799)
gitano X male			-0.526		-0.648			-0.689		-0.819
			(0.663)		(0.983)			(0.663)		(0.996)
mixed X male				-1.543**	-1.725*				-1.491**	-1.685*
				(0.641)	(0.971)				(0.646)	(0.987)
gitano X mixed X male					0.613					0.650
					(1.336)					(1.353)
age						0.011	0.010	0.012	0.009	0.010
						(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
hincome						-0.045	-0.042	-0.056	-0.045	-0.056
						(0.135)	(0.135)	(0.133)	(0.129)	(0.130)
Constant	6.380***	6.224***	6.297***	6.087***	5.941***	5.928***	5.801***	5.793***	5.732***	5.575***
	(0.262)	(0.296)	(0.278)	(0.282)	(0.322)	(0.836)	(0.840)	(0.846)	(0.825)	(0.839)
F	0.870	1.042	0.780	2.296*	1.420	0.927	1.000	0.969	1.816*	1.347
Log-likelihood	-754.844	-754.136	-754.510	-751.791	-751.350	-753.927	-753.314	-753.368	-751.077	-750.513
R^2	0.009	0.013	0.011	0.028	0.030	0.014	0.018	0.018	0.032	0.036
Obs.	313	313	313	313	313	313	313	313	313	313

Notes: OLS estimates. Robust standard errors are shown in parentheses. Dependent variable: contribution (euros; range 0–10). Main explanatory variables: *gitano*, *mixed*, *male* and their interactions are binary variables (0/1). Columns 1b–5b repeat the regressions adding control variables: *age* (range 16–-82), *hincome* (household income: range 0–9, from "0 euros/month" to "more than 5,000 euros/month"; 12 missing values were imputed using OLS regression with *gitano*, *age*, and *male* as explanatory variables). * $P < .10$, ** $P < .05$, *** $P < .01$.

Table S2. *The determinants of punishment (aggregate)*

depvar: *punishment*	(1a)	(2a)	(3a)	(4a)	(5a)	(1b)	(2b)	(3b)	(4b)	(5b)
gitano	-0.362***	-0.870***	-0.616***	-0.459***	-1.050***	-0.304**	-0.813***	-0.559***	-0.382**	-0.962***
	(0.116)	(0.156)	(0.122)	(0.114)	(0.173)	(0.153)	(0.162)	(0.155)	(0.149)	(0.184)
mixed	-0.065	-0.418**	-0.080	-0.492***	-0.748***	-0.052	-0.403**	-0.075	-0.477***	-0.738***
	(0.148)	(0.193)	(0.145)	(0.158)	(0.221)	(0.152)	(0.193)	(0.150)	(0.159)	(0.223)
male	0.134	0.120	-0.147	-0.373***	-0.656***	0.126	0.119	-0.166	-0.386***	-0.668***
	(0.118)	(0.124)	(0.189)	(0.136)	(0.186)	(0.123)	(0.130)	(0.198)	(0.142)	(0.193)
differ	0.074***	0.076***	0.073***	0.078***	0.078***	0.074***	0.076***	0.073***	0.077***	0.078***
	(0.017)	(0.017)	(0.017)	(0.016)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
meancont2others	0.077**	0.070**	0.080***	0.073**	0.073**	0.077***	0.070**	0.080***	0.074***	0.073***
	(0.030)	(0.030)	(0.030)	(0.029)	(0.028)	(0.030)	(0.029)	(0.029)	(0.028)	(0.028)
gitano X mixed		0.807***			0.730***		0.814***			0.738***
		(0.228)			(0.242)		(0.230)			(0.244)
gitano X male			0.574**		0.700***			0.579**		0.690***
			(0.231)		(0.241)			(0.233)		(0.249)
mixed X male				1.085***	1.165***				1.113***	1.192***
				(0.223)	(0.358)				(0.224)	(0.355)
gitano X mixed X male					-0.476					-0.488
					(0.434)					(0.438)
age						0.002	0.003	0.001	0.004	0.004
						(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
hincome						0.019	0.014	0.027	0.016	0.024
						(0.043)	(0.044)	(0.044)	(0.044)	(0.045)
Constant	0.396*	0.652***	0.474**	0.636***	0.848***	0.224	0.458	0.326	0.366	0.585*
	(0.234)	(0.246)	(0.230)	(0.224)	(0.239)	(0.325)	(0.318)	(0.320)	(0.319)	(0.320)
Chi^2	38.045***	52.992***	53.771***	58.644***	78.609***	40.733***	55.917***	54.738***	64.577***	80.230***
Log-likelihood	-1712.275	-1707.827	-1709.767	-1703.265	-1698.768	-1712.103	-1707.575	-1709.597	-1702.715	-1698.316
R^2	0.048	0.066	0.059	0.082	0.095	0.049	0.067	0.059	0.084	0.097
Obs.	939	939	939	939	939	939	939	939	939	939

Notes: GLMM random effects estimates. Dependent variable: punishment (euros reduced per target; range 0–9). Main explanatory variables: same as in Table S1 + *differ* (punisher's contribution – target's contribution, from -10 to 10) + *meancont2others* (mean contribution of other 2 group members, range 0–10). See notes in Table S1.

Table S3. *The determinants of punishment (mixed groups)*

depvar: *punishment*	(1a)	(2a)	(3a)	(4a)	(5a)	(1b)	(2b)	(3b)	(4b)	(5b)
gitano	-0.140	-0.312	-0.141	-0.139	-0.356*	-0.103	-0.275	-0.101	-0.104	-0.317
	(0.173)	(0.219)	(0.173)	(0.174)	(0.216)	(0.240)	(0.270)	(0.240)	(0.243)	(0.271)
male	0.579***	0.579***	0.572***	0.583***	0.572***	0.656***	0.656***	0.648***	0.664***	0.651***
	(0.179)	(0.179)	(0.177)	(0.178)	(0.175)	(0.181)	(0.181)	(0.179)	(0.180)	(0.178)
targetgit	0.103	-0.072	0.102	0.098	-0.103	0.104	-0.071	0.103	0.099	-0.103
	(0.151)	(0.258)	(0.151)	(0.150)	(0.256)	(0.151)	(0.258)	(0.151)	(0.150)	(0.255)
differ	0.096***	0.096***	0.084**	0.044*	0.030	0.094***	0.094***	0.081**	0.041*	0.024
	(0.022)	(0.022)	(0.034)	(0.023)	(0.046)	(0.022)	(0.022)	(0.034)	(0.022)	(0.044)
meancont2others	0.051	0.052	0.054	0.048	0.052	0.054	0.054	0.056	0.050	0.054
	(0.042)	(0.042)	(0.044)	(0.038)	(0.040)	(0.040)	(0.040)	(0.042)	(0.037)	(0.038)
gitano X targetgit		0.342			0.411		0.342			0.411
		(0.321)			(0.322)		(0.321)			(0.322)
gitano X differ			0.024		0.018			0.024		0.024
			(0.045)		(0.052)			(0.043)		(0.052)
targetgit X differ				0.105***	0.091**				0.106***	0.096**
				(0.037)	(0.046)				(0.036)	(0.046)
gitano X targetgit X differ					0.057					0.049
					(0.074)					(0.078)
age						0.006	0.006	0.006	0.006	0.006
						(0.007)	(0.007)	(0.007)	(0.007)	(0.008)
hincome						-0.029	-0.029	-0.028	-0.031	-0.030
						(0.055)	(0.055)	(0.055)	(0.055)	(0.054)
Constant	0.131	0.246	0.127	0.127	0.257	-0.087	0.028	-0.098	-0.091	0.037
	(0.305)	(0.330)	(0.310)	(0.290)	(0.319)	(0.494)	(0.509)	(0.497)	(0.508)	(0.522)
Chi^2	29.914***	34.469***	33.038***	39.045***	46.292***	31.538***	37.257***	34.802***	43.635***	53.476***
Log-likelihood	-857.328	-856.552	-857.163	-853.233	-850.976	-856.623	-855.847	-856.455	-852.468	-850.243
R^2	0.081	0.083	0.080	0.095	0.102	0.091	0.093	0.091	0.106	0.113
Obs.	465	465	465	465	465	465	465	465	465	465

Notes: GLMM random effects estimates. Dependent variable: punishment (euros reduced per target; range 0–9). Main explanatory variables: same as in Table S2 + *targetgit* (binary variable: whether the target is gitano (0/1)). See notes in Table S1.