Third-party punishment is common across human societies but varies with socioecology

Zachary H. Garfield[1†*], Grégory Fiorio[2], Ana Kaneko Ebert[3], Sophia Iben El Rhazi[1], Francesca De Petrillo[4†], Catherine Molho[5†]

[1]Africa Institute for Research in Economics and Social Sciences, University Mohammed VI Polytechnic, Rabat, Morocco; [2]Institut Jean Nicod, Département d'études cognitives, École normale supérieure, Université PSL, EHESS, CNRS, Paris, France; [3] Vrije Universiteit, Amsterdam, The Netherlands; [4]School of Psychology & Biosciences Institute, Newcastle University, Newcastle upon Tyne, England; [5]Department of Social and Behavioral Sciences, Toulouse School of Economics, Toulouse, France.

*Corresponding author, zachary.garfield@um6p.ma [†] Senior author

Abstract

Punishment systems are central to maintaining human cooperation, yet how they vary across societies remains poorly understood. Different theoretical frameworks make distinct predictions about who third-party punishment (3PP) defends and the socioecological conditions under which it occurs. We test such predictions using a new cross-cultural dataset of 3961 ethnographic texts (coded 11893 times) from 60 largely nonindustrial societies. We code evidence for second- and third-party punishment and, when third-party punishment is documented, whether it is enacted on behalf of kin, social partners, or the broader community. Using Bayesian multilevel regression models, we test preregistered hypotheses linking these different forms of punishment to society-level measures of complexity (i.e., technological and social differentiation, resource-use intensification, community size) and kinship intensity. We find that third-party punishment is widespread across societies and ethnographic texts, occurring more frequently than second-party punishment. Aggregated across relational subtypes, third-party punishment shows weak associations with socioecology; disaggregated, it shows distinct and theoretically meaningful patterns. Generalized 3PP on behalf of community members is positively associated with technological and social differentiation, while kin-directed 3PP is positively associated with kinship intensity and negatively associated with technological and social differentiation. Contrary to our prediction, partner-directed 3PP shows a robust negative association with resource-use intensification. Second-party punishment is negatively associated with technological and social differentiation. These findings show that socioecological structures shape not only the presence of punishment but also the forms it takes and whom it is used to defend, revealing distinct punishment regimes rather than a single enforcement continuum.

Main

Unlike other social mammals, humans often cooperate beyond their close kin and social partners in large groups of unrelated individuals[1]. Such large-scale cooperation yields substantial benefits. Cooperative groups can hunt larger game[2,3], preserve common-pool resources[4], and wage war to acquire territory and livestock[5–7]. Yet large-scale cooperation is perpetually threatened by those who violate the norms and institutions that sustain it. Societies therefore rely on punishment to deter such defection[8–11] and restore social cohesion[12–14]. Accordingly, most scholars agree that punishment plays a central role in sustaining and shaping human cooperation[9,15]. Others, however, have questioned the universality of this function[11,16], pointing out that punishment is neither costless nor a panacea: it requires time and resources, invites counter-punishment[17,18], and can erode group cooperation when used antisocially[19]. Understanding when punishment stabilizes cooperation therefore requires explaining not only whether punishment occurs, but also what form it takes and who is willing to enforce it.

Substantial research has demonstrated that punishment is widespread across human societies, but highly variable in form and function[13,20–24]. People punish through a wide range of direct and indirect tactics, including corporal sanctions, material penalties, social exclusion, gossip, ridicule, and public condemnation[20]. To capture this diversity, we adopt a broad definition of punishment as any action that imposes costs on individuals in response to perceived offenses or norm violations. This inclusive definition allows us to capture the full range of sanctioning behaviors documented in the ethnographic record, beyond formal legal or institutional penalties.

Importantly, people do not only punish those who harm them but also punish on behalf of others[25,26]. Yet, the prevalence of such third-party punishment in smaller-scale societies remains debated[16,27,28]. Moreover, the relational context of punishment–whether enacted on behalf of the victim, their kin, their social partners, or the broader community–remains underexplored. Most experimental work has focused on interactions among strangers[19,26,29–31], obscuring the relational contexts that shape real-world sanctioning and overlooking evidence that moral decision-making is fundamentally shaped by social relationships[32]. Some forms, such as second-party punishment or kin-based third-party punishment, have been characterized as “vengeful” or self-interested[27]. Other forms, like third-party punishment on behalf of community members, are often viewed as “prosocial”[15,26,33]. Distinguishing among these forms also allows us to test hypotheses from kin selection, reciprocal altruism, and strong reciprocity theories[34–36], each of which makes specific predictions about when punishment occurs and whom it serves to defend.

To systematically examine variation in punishment, we distinguish two broad punishment types: second-party punishment (2PP), in which victims retaliate directly against offenders, and third-party punishment (3PP), where individuals not directly involved in the offense intervene to punish on behalf of others. Within 3PP, we further differentiate punishment types based on the relationship between punisher and victim: kin-based 3PP, in which individuals punish to defend relatives; partner-based 3PP, enacted on behalf of

friends or allies; and generalized 3PP, which involves punishing on behalf of community members or group norms in the absence of strong personal ties. These distinctions are key in ongoing debates on whether third-party punishment was a strong selective force in human cooperation[9,15] or a relatively rare phenomenon restricted to specific socioecological contexts[16,37]. Advancing these debates requires comparative data that encompasses the full spectrum of human socioecological diversity. By examining rich descriptions of punishment in politically autonomous, nonindustrial societies, we can assess how features of socioecology shape its occurrence across different relational contexts, addressing the question of who punishes on whose behalf.

2PP is commonly explained by deterrence: by imposing costs on offenders, victims signal that exploiting them is costly, discouraging further harm by the offender (specific deterrence) or others who observe the punishment (general deterrence)[38,39]. 2PP does not require intermediaries and is often driven by emotional responses such as anger or discontent[40–43], which serve adaptive functions in facilitating social support and discouraging exploitation[44]. In small-scale, politically autonomous societies where formal institutions are limited, it may serve as a primary mechanism for conflict regulation[13,27,28]. Whilst 2PP often reflects personal grievance and self-interested motives, 3PP often serves broader social functions, including signaling norm adherence, enhancing reputation[29,31], and sustaining impartial enforcement in larger, more complex groups[15,26,45].

Kin-based third-party punishment (3PP-kin) is often explained by kin selection theory, the evolutionary principle that individuals are more likely to support and protect biological relatives due to shared genetic interests[35]. From this perspective, punishing those who harm kin can enhance inclusive fitness by defending genetically related individuals[39,46,47]. These behaviors are likely supported by proximate psychological mechanisms, such as heightened empathy, obligation, or interdependence, that prioritize kin relationships[48,49]. Although kin selection and fitness interdependence provide universal motives for 3PP-kin, we expect these tendencies to be amplified in kinship-intensive societies for two reasons. First, individuals in such societies interact and rely more heavily on extended kin networks, raising the direct benefits of protecting family members. Second, kinship-intensive societies tend to institutionalize stronger obligations to defend relatives, with greater social repercussions for failing to do so. Consistent with this, tight kinship systems are associated across societies with communal moral values and relational forms of norm enforcement, including revenge, punishment, and public shame, rather than universalistic norms or impartial sanctioning[50,51].

Third-party punishment on behalf of social partners (3PP-partner) may arise through reciprocal altruism, in which individuals help non-relatives when cooperation is expected to be reciprocated over repeated interactions[36], or through interdependence, where individuals' outcomes are closely tied to those of difficult-to-replace partners[52,53]. Partner-based punishment thus functions to sustain the valuable social ties on which cooperation depends. Punishing transgressions against social partners is especially adaptive in contexts where cooperation depends on durable alliances beyond close kin, such as in societies with larger community size, greater resource-use intensification, or increased

technological and social differentiation. In these settings, individuals rely on exchange partners, coalitional allies, or interdependent specialists for access to resources, protection, or information.

Generalized third-party punishment (3PP-general) has been explained through multiple theoretical frameworks. Some accounts emphasize individual-level benefits, suggesting that third-party punishment can serve as a costly signal of trustworthiness or a means of enhancing personal reputation[31,54]. Others highlight group-level functions, such as norm adherence and strong reciprocity, whereby individuals are predisposed to enforce social norms even in the absence of direct benefit, helping to stabilize cooperation in large, impersonal groups[15,25,34]. These mechanisms are expected to become more salient in societies characterized by greater technological and social differentiation and larger community size, where interactions commonly extend beyond close kin and close social partners[21,27,55]. As social life becomes more interdependent and specialized, through, for example, labor differentiation, political integration, and urbanization, norm enforcement cannot rely solely on kinship or direct reciprocity[15,26]. Maintaining cooperation requires more generalized and impartial systems of sanctioning. Institutionalized punishment systems, such as formal legal codes or structured communal sanctions, often emerge under these conditions[56,57]. Taken together, these theoretical perspectives imply that punishment should vary systematically with socioecology: the social, political, and economic features of societies should shape not only whether punishment occurs, but who enforces it and on whose behalf. Indeed, prior cross-cultural work has documented associations between socioecological conditions and punishment systems. For instance, larger and more socially differentiated societies tend to exhibit greater reliance on institutional enforcement and impersonal third-party punishment (including capital punishment[20]) while societies with more intensive kinship systems often emphasize relationally embedded sanctions such as retaliation, shame, or collective responsibility[22,50,51].

However, existing approaches face several limitations. First, much experimental work has examined punishment using abstract economic games among anonymous strangers[10,19,25,26,34], stripping away the relational ties and institutional contexts that shape real-world sanctioning. These studies provide behavioral evidence of sanctioning preferences, but their ecological validity is limited. Second, vignette-based studies often employ more contextually grounded scenarios but rely on self-reported judgments about how participants would punish, introducing concerns about social desirability and hypothetical bias[48–50,58,59]. Third, existing ethnographic case studies and comparisons have typically been based on relatively small samples of societies[13,28,60], limiting their ability to estimate cross-cultural prevalence or systematically test socioecological predictors. As a result, we still lack comparative evidence that can estimate how common different punishment forms are across societies and test whether distinct socioecological conditions predict distinct punishment types.

To address this gap, we constructed a novel database of punishment by coding primary ethnographic descriptions from a stratified random sample of culturally diverse, largely

nonindustrial societies. Our database distinguishes punishment by its relational context (i.e., kin, partners, or community members), thereby addressing key limitations of prior work. Rather than relying on abstract economic games or self-reported vignette responses, our data incorporate two distinct forms of ethnographic evidence: accounts of actual sanctioning events embedded in real social relationships and institutional contexts, and cultural models articulating normative expectations about how punishment should operate within a given society. With 5478 ethnographic paragraphs coded 11893 times by 8 independent coders across 60 societies, our database enables large-scale, paragraph-level assessment of punishment diversity. These data allow us to model punishment in its relational and institutional contexts and link its diverse forms to theoretically grounded socioecological dimensions via confirmatory analyses and multilevel modeling.

We first report descriptive statistics on the frequency of second- and third-party punishment in our ethnographic sample, including how often each relational form of third-party punishment appears across societies. We then test preregistered hypotheses linking socioecological variation to the presence and form of punishment, examining how kinship intensity and societal differentiation predict distinct punishment types. Through this framework, we empirically evaluate competing theoretical accounts (i.e., kin selection, reciprocal altruism, and strong reciprocity) of how human groups regulate cooperation under different ecological and institutional conditions. We additionally conduct exploratory analyses examining the cultural domains associated with each punishment type using subject codes provided by the ethnographic source database.

To test how socioecological conditions shape punishment, we focus on four dimensions: technological and social differentiation (TSD), resource-use intensification (RI), kinship intensity (KII), and community size (CS). Rather than treating societal "complexity" as a unitary construct, we deconstruct it into distinct components that evolutionary theories link to the presence and form of punishment in different ways. TSD and RI are derived from latent factor models applied to ethnographic data by Ringen et al.[61] and capture the specialization of labor and political integration (TSD) and investment in intensified production systems (RI)(Fig. 7). KII measures the strength of kin-based organization in structuring social and economic life[62] and CS indexes typical settlement and political unit scale[21,55]. We derive specific predictions linking each dimension to distinct punishment types in turn.

We expect second-party punishment to be more prevalent in societies with intensive kinship structures, where social life is organized around known, interdependent kin, making direct retaliation both feasible and socially expected, and where formal third-party institutions are less developed[50]. In kin-intensive societies, the costs of failing to defend oneself against transgressions are amplified both materially–because one's well-being is tightly linked to one's standing within kin networks–and reputationally, because tolerating offenses may signal weakness to kin and non-kin alike in the absence of reliable third-party enforcement. Conversely, we expect second-party punishment to decline with increasing societal scale and intensification, including larger community size, resource-use

intensification, and technological and social differentiation, as collective action problems create demand for third-party enforcement[21,26].

For kin-based third-party punishment, we predict positive associations with kinship intensity through two complementary pathways. First, a direct benefits pathway: in societies with intensive kinship structures, individuals interact more heavily with extended kin and derive greater material and social benefits from these relationships, raising the direct stakes of protecting family members from harm. Second, a norm-obligation pathway: kinship-intensive societies tend to institutionalize stronger expectations to defend relatives, with greater reputational costs for failing to do so–creating social pressure to punish on behalf of kin independent of any direct material gain.

Partner-based third-party punishment should be more common in societies characterized by greater technological and social differentiation and larger community size, where individuals increasingly rely on non-kin allies, exchange partners, and interdependent specialists. In these settings, the value of specific partnerships rises because partners provide access to resources, skills, or protection that cannot easily be obtained elsewhere. Punishing transgressions against such partners is therefore adaptive. It signals loyalty, deters future exploitation of valued allies, and helps sustain the cooperative relationships on which individuals depend. This logic follows from reciprocal altruism and fitness interdependence accounts, which predict that individuals will invest in protecting relationships when those relationships are durable, valuable, and difficult to replace[36,52,53].

Finally, we predict that generalized third-party punishment will be positively associated with all three measures of societal scale and complexity: community size, resource-use intensification, and technological and social differentiation. These conditions favor impartial enforcement systems over kin-based or relational sanctioning, as cooperation in large, impersonal groups cannot be sustained through kinship or direct reciprocity alone[15,26].

Results

We compiled a large, systematic sample of ethnographic excerpts relevant to punishment from the electronic Human Relations Area Files World Cultures database. This sample comprises 15380 passages from 843 primary documents spanning 60 culturally diverse societies[63]. We then developed and refined a coding scheme to identify whether passages contained evidence of a social offense and, when an offense was present, whether they provided evidence for distinct punishment types. A random subset of these passages was independently coded by at least two coders (n = 3961). We coded evidence for second-party punishment and third-party punishment and, when third-party punishment was present, further classified it as occurring on behalf of kin, social partners, or as generalized third-party punishment enacted by other community members or institutional systems. We also coded unspecified punishment (i.e., clear evidence of punishment for which second- versus third-party enforcement could not be determined) (Table 3). The 8 coders each evaluated a variable number of unique paragraphs (mean = 1115.9, SD = 1130, range

= 25–3477). All inferential models include varying intercepts for coder, paragraph, document, and society, thereby incorporating coder-level disagreement and heterogeneity directly into parameter uncertainty rather than treating codings as independent observations. The full coding protocol and coder training procedures are provided in the Methods and Supplementary Materials.

Descriptive results

Assessing the evidence for second and third-party punishment

At the paragraph level, coders identified evidence for 2PP in 2.9% (80% HDI 2.1%–4.5%; 90% HDI 1.9%–5.5%) of coded passages and evidence for 3PP in 36.9% (80% HDI 30.1%–44.7%; 90% HDI 28.1%–47.3%), with 6.0% (80% HDI 4.1%–9.5%; 90% HDI 3.7%–11.4%) classified as unspecified punishment (Fig. 1A). When 3PP was present, it was most often coded as 3PP-general (25.1% (80% HDI 19.1%–32.8%; 90% HDI 17.6%–35.7%)), followed by 3PP-partner (6.2% (80% HDI 1.3%–25.1%; 90% HDI 0.9%–33.6%)) and 3PP-kin (0.8% (80% HDI 0.5%–2.4%; 90% HDI 0.4%–5.1%)) (Fig. 1B).

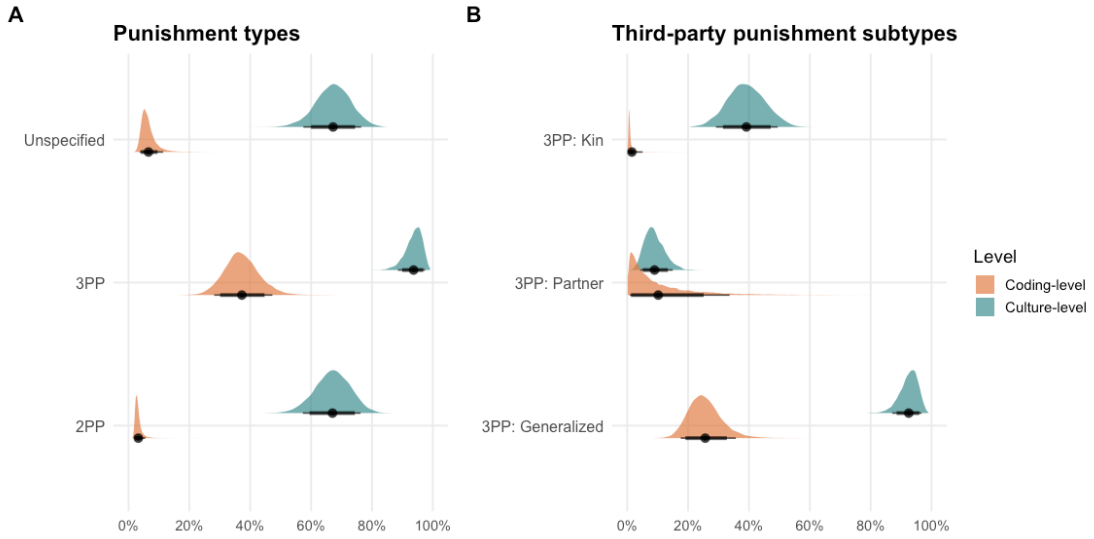

Figure 1: Coding-level and culture-level posterior distributions for punishment types and third-party punishment subtypes. Points show posterior means with 80% and 90% credible intervals.

To characterize cross-societal prevalence, we constructed society-level presence indicators. For each punishment type, a society was coded as exhibiting that type if at least two paragraphs from that society contained confirmed supporting evidence (i.e., each met the multi-coder confirmation criterion of two independent coders recording 'evidence for'). Under this definition, 2PP was present in 67.2% (80% HDI 59.5%–74.4%; 90% HDI 57.3%–76.2%) of societies and 3PP was present in 94.1% (80% HDI 89.9%–96.8%; 90% HDI 88.5%–97.4%), with unspecified punishment present in 67.3% (80% HDI 60.0%–74.4%; 90% HDI 57.4%–76.5%) (Fig. 1A). Among societies exhibiting 3PP, 3PP-general was most common (92.9% (80% HDI 88.5%–96.0%; 90% HDI 87.1%–96.7%)), followed by 3PP-kin

(39.0% (80% HDI 31.5%–47.2%; 90% HDI 29.2%–49.5%)) and 3PP-partner (8.6% (80% HDI 5.1%–13.4%; 90% HDI 4.3%–15.0%)) forms (Fig. 1B). Partner-directed 3PP was the rarest subtype, with fewer than 10% of societies showing two or more confirmed paragraphs of partner-based enforcement. We discuss possible interpretations of this rarity, including measurement-validity concerns, in the Discussion.

Testing hypotheses on evidence for second and third-party punishment and cultural variation

We next evaluated preregistered hypotheses linking socioecological variation to punishment types using Bayesian multilevel logistic regression models. These models predict coder–paragraph-level evidence for punishment from society-level predictors (RI, TSD, CS, and KII), all z-scored within each outcome before estimation. Across outcomes, we report posterior medians (log-odds), 80% and 90% credible intervals (HDIs), and the posterior probability of direction (pd), defined as the proportion of posterior draws sharing the sign of the median. For each punishment type, we first highlight associations that were consistent across univariate (single predictor only) and multivariate (all four predictors included) models, and then describe how estimates changed when predictors were modeled jointly. The status of each preregistered prediction is summarized in Table 1. Detailed coefficients for all preregistered and exploratory predictor/outcome combinations are provided in Table 2 following the per-outcome subsections below.

Table 1: Preregistered hypothesis scorecard. Predicted: direction predicted in preregistration (https://osf.io/79g4e). Observed: direction of multivariate posterior median. Support: classification based on multivariate model. 'Supported' = 90% HDI excludes 0 in predicted direction; 'Moderately supported' = 80% HDI excludes 0 in predicted direction; 'Suggestive (NS)' = posterior median in predicted direction with pd > 0.80 but 80% HDI includes 0; 'Not supported' = posterior median against predicted direction or pd

Outcome	Predictor	Predicted	Observed	β [90% HDI]	Support
2PP	RI	↓	↑	0.04 [-0.23, 0.33]	Not supported
	TSD	↓	↓	-0.32 [-0.57, -0.07]	Supported
	CS	↓	↓	-0.02 [-0.20, 0.16]	Not supported
	KII	↑	↓	-0.02 [-0.18, 0.15]	Not supported
3PP (aggregate)	RI	↑	↓	-0.03 [-0.27, 0.20]	Not supported
	TSD	↑	↑	0.13 [-0.06, 0.35]	Suggestive (NS)
	CS	↑	↑	0.01 [-0.14, 0.15]	Not supported
	KII	↓	↑	0.04 [-0.10, 0.17]	Not supported
3PP-General	RI	↑	↑	0.11 [-0.16, 0.36]	Not supported
	TSD	↑	↑	0.18 [-0.03, 0.42]	Moderately supported

Table 1: Preregistered hypothesis scorecard. Predicted: direction predicted in preregistration (https://osf.io/79g4e). Observed: direction of multivariate posterior median. Support: classification based on multivariate model. 'Supported' = 90% HDI excludes 0 in predicted direction; 'Moderately supported' = 80% HDI excludes 0 in predicted direction; 'Suggestive (NS)' = posterior median in predicted direction with pd > 0.80 but 80% HDI includes 0; 'Not supported' = posterior median against predicted direction or pd

Outcome	Predictor	Predicted	Observed	β [90% HDI]	Support
	CS	↑	↓	-0.02 [-0.17, 0.15]	Not supported
3PP-Partner	RI	↑	↓	-0.67 [-1.14, -0.19]	Not supported
	TSD	↑	↑	0.22 [-0.21, 0.65]	Suggestive (NS)
	CS	↑	↑	0.18 [-0.13, 0.50]	Suggestive (NS)
3PP-Kin	KII	↑	↑	0.26 [-0.03, 0.55]	Moderately supported

Second-party punishment

Figure 2A shows associations between sociocultural predictors and 2PP. The most robust pattern was a negative association between technological and social differentiation (TSD) and 2PP, observed in both the univariate model (median β = -0.29 (80% HDI -0.40–-0.19; 90% HDI -0.42–-0.17; pd = 1.00)) and the multivariate model (median β = -0.32 (80% HDI -0.52–-0.12; 90% HDI -0.57–-0.07; pd = 0.98)). The agreement between specifications indicates that TSD's negative association with 2PP holds both as a total association and as a direct association after partialing out the other predictors.

Resource-use intensification (RI) and community size (CS) were also negatively associated with 2PP in univariate models (median β = -0.25 (80% HDI -0.36–-0.15; 90% HDI -0.39–-0.12; pd = 1.00); median β = -0.18 (80% HDI -0.29–-0.07; 90% HDI -0.32–-0.04; pd = 0.98)), but these direct associations attenuated substantially in the multivariate specification (median β = 0.04 (80% HDI -0.17–0.26; 90% HDI -0.23–0.33; pd = 0.59); median β = -0.02 (80% HDI -0.16–0.12; 90% HDI -0.20–0.16; pd = 0.56)). Following the framework outlined in the Methods, this is consistent with RI and CS exerting their associations with 2PP largely through TSD-mediated pathways rather than through direct effects.

Contrary to our preregistered prediction that kinship intensity (KII) would positively predict 2PP, we found no evidence for this association in either the univariate (median β = -0.07 (80% HDI -0.20–0.06; 90% HDI -0.24–0.09; pd = 0.77)) or multivariate model (median β = -0.02 (80% HDI -0.14–0.11; 90% HDI -0.18–0.15; pd = 0.57)).

Together, these findings partially support the hypothesis that sociocultural complexity is associated with reduced second-party punishment, with TSD showing the clearest direct association and RI and CS contributing primarily through TSD-mediated pathways.

Third-party punishment

For third-party punishment as an aggregated category, associations with all four predictors were weak in both univariate and multivariate specifications (Fig. 2B; full coefficients in Table 2). RI and TSD showed modest positive univariate associations (median β = 0.10 (80% HDI 0.02–0.19; 90% HDI -0.00–0.21; pd = 0.95); median β = 0.12 (80% HDI 0.04–0.20; 90% HDI 0.02–0.22; pd = 0.97)) that attenuated under multivariate specification, while CS and KII showed weak associations across both specifications. This aggregated analysis treats all forms of third-party punishment as equivalent and obscures substantively important heterogeneity across relational subtypes. As we show in the analyses that follow, the apparent weakness of socioecological associations at the aggregate level masks robust and theoretically distinct associations with kin-directed, partner-directed, and generalized 3PP. We therefore interpret the aggregated 3PP analysis as a check that no single dimension of socioecology drives third-party punishment broadly, and turn to the subtype analyses for the substantive findings.

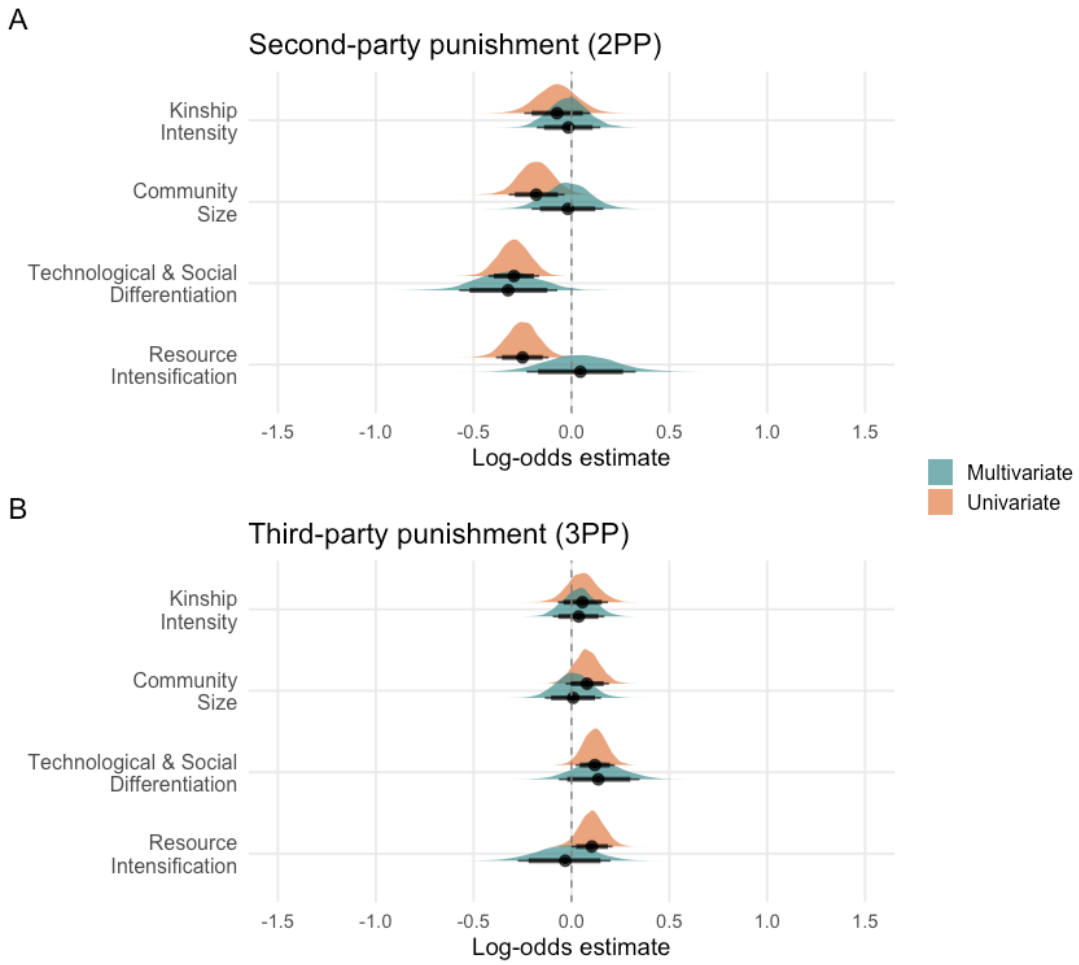

Figure 2: Predictors of second- and third-party punishment. Points show posterior means with 80% and 90% credible intervals. Orange distributions = univariate model (single

predictor); teal distributions = multivariate model (all four predictors included simultaneously).

Generalized third-party punishment

Figure 3A shows associations between socioecological factors and generalized third-party punishment. In the univariate models we observe evidence in support of our preregistered hypotheses, according to which sociocultural indicators of "complexity" would be associated with a stronger presence of generalized third-party punishment. Resource-use intensification (median β = 0.27 (80% HDI 0.18–0.36; 90% HDI 0.15–0.39; pd = 1.00)), technological and social differentiation (median β = 0.27 (80% HDI 0.17–0.35; 90% HDI 0.15–0.38; pd = 1.00)), and community size (median β = 0.17 (80% HDI 0.07–0.27; 90% HDI 0.04–0.30; pd = 0.98)) were all positively associated with generalized third-party punishment.

In the multivariate model, which estimates each predictor's direct association after partialing out the others, technological and social differentiation retained a positive association (median β = 0.18 (80% HDI 0.01–0.37; 90% HDI -0.03–0.42; pd = 0.92)), while resource-use intensification (median β = 0.11 (80% HDI -0.10–0.30; 90% HDI -0.16–0.36; pd = 0.75)) and community size (median β = -0.02 (80% HDI -0.14–0.12; 90% HDI -0.17–0.15; pd = 0.56)) attenuated substantially. Given Ringen et al.'s evidence that resource-use intensification is causally upstream of technological and social differentiation, the divergence between RI's univariate and multivariate coefficients is consistent with RI's effect on generalized third-party punishment operating largely through its effect on TSD, rather than through a direct pathway. In other words, intensified production systems appear to facilitate generalized 3PP primarily by promoting the specialized roles, formal procedures, and political integration that constitute TSD, which in turn support institutionalized norm enforcement.

We did not make a preregistered prediction regarding the association between kinship intensity and generalized third-party punishment, and results indicate that this relationship is weak or non-existent in both models (median β = 0.02 (80% HDI -0.10–0.14; 90% HDI -0.14–0.18; pd = 0.58); median β = -0.02 (80% HDI -0.13–0.10; 90% HDI -0.16–0.13; pd = 0.57)). Together, these findings support the hypothesis that generalized third-party punishment is facilitated by sociocultural complexity, with technological and social differentiation showing the most robust direct association and resource-use intensification contributing primarily through TSD-mediated pathways.

Third-party punishment on behalf of social partners

Figure 3B shows associations between socioecological factors and partner-directed third-party punishment. Contrary to our preregistered prediction that sociocultural indicators of "complexity" would be positively associated with partner-directed third-party punishment, resource-use intensification was robustly negatively associated with 3PP-partner in the univariate model (median β = -0.36 (80% HDI -0.53–-0.18; 90% HDI -0.59–-0.13; pd = 1.00)), and this negative association strengthened in the multivariate specification (median β = -

0.67 (80% HDI -1.04–-0.31; 90% HDI -1.14–-0.19; pd = 0.99)). Under the framework outlined in the Methods, the univariate estimate captures RI's total association with 3PP-partner, while the multivariate estimate captures RI's direct association after partialing out the TSD-mediated pathway. The strengthening of RI's negative coefficient when TSD is included is consistent with RI exerting a direct negative effect on partner-based enforcement that is partially offset, in the univariate specification, by a small positive contribution running through TSD. Substantively, this suggests that resource-use intensification reduces reliance on ally-based enforcement, even as the technological and social differentiation it tends to coevolve with may modestly support such enforcement.

Technological and social differentiation showed a negative univariate association with 3PP-partner (median β = -0.23 (80% HDI -0.41–-0.05; 90% HDI -0.46–0.00; pd = 0.95)) that weakened and reversed sign in the multivariate specification (median β = 0.22 (80% HDI -0.10–0.55; 90% HDI -0.21–0.65; pd = 0.81)). Community size showed a similar pattern (median β = -0.10 (80% HDI -0.30–0.09; 90% HDI -0.35–0.14; pd = 0.76); median β = 0.18 (80% HDI -0.06–0.42; 90% HDI -0.13–0.50; pd = 0.84)). These reversals indicate that TSD's and CS's apparent negative associations in univariate specifications are driven by their correlation with RI; once RI's direct negative effect is partialed out, TSD and CS show weak positive direct associations with partner-directed enforcement, though both posterior intervals span zero.

Kinship intensity showed a weak and inconsistent positive association with partner-directed punishment across specifications (median β = 0.13 (80% HDI -0.09–0.33; 90% HDI -0.15–0.40; pd = 0.78); median β = 0.10 (80% HDI -0.12–0.31; 90% HDI -0.17–0.37; pd = 0.72)). Notably, this contrasts with the stronger and more consistent positive association between kinship intensity and kin-directed third-party punishment (see Fig. 3C), suggesting that kin-based enforcement is more tightly anchored in intensive kinship systems than punishment enacted on behalf of social partners.

Together, these findings indicate that our preregistered prediction for 3PP-partner was not supported. Rather than increasing with sociocultural complexity, partner-directed punishment shows a robust negative association with resource-use intensification specifically. We discuss possible interpretations of this unexpected pattern in the Discussion.

Third-party punishment on behalf of kin

Figure 3C shows associations between socioecological factors and kin-directed third-party punishment. Our only preregistered hypothesis for this outcome concerned a positive association between kinship intensity and kin-directed third-party punishment. Consistent with this prediction, kinship intensity showed a positive association in the univariate model (median β = 0.19 (80% HDI -0.04–0.42; 90% HDI -0.10–0.50; pd = 0.86)) that strengthened in the multivariate specification (median β = 0.26 (80% HDI 0.04–0.49; 90% HDI -0.03–0.55; pd = 0.93)), with the 90% credible interval narrowly crossing zero.

Associations with resource-use intensification and technological and social differentiation were not preregistered and should therefore be considered exploratory. In univariate models, both RI and TSD showed robust negative associations with 3PP-kin (median β = -0.37 (80% HDI -0.56–-0.17; 90% HDI -0.61–-0.12; pd = 0.99); median β = -0.40 (80% HDI -0.58–-0.21; 90% HDI -0.63–-0.16; pd = 1.00)). In the multivariate specification, RI's coefficient attenuated substantially (median β = -0.07 (80% HDI -0.46–0.32; 90% HDI -0.58–0.43; pd = 0.59)), while TSD retained a negative association of similar magnitude though with greater uncertainty (median β = -0.40 (80% HDI -0.75–-0.04; 90% HDI -0.85–0.06; pd = 0.92)). Following the framework outlined in the Methods, this divergence is consistent with RI's negative association with kin-directed punishment operating largely through its effect on TSD: once TSD's direct negative association is partialed out, little independent contribution from RI remains. The substantive implication is that resource-use intensification appears to reduce kin-directed enforcement primarily by promoting the technological and social differentiation that displaces lineage-based authority, rather than through a direct pathway. Community size showed little consistent association in either specification (median β = -0.19 (80% HDI -0.40–0.00; 90% HDI -0.45–0.06; pd = 0.90); median β = 0.03 (80% HDI -0.22–0.29; 90% HDI -0.30–0.37; pd = 0.56)).

Together, these findings support the preregistered hypothesis that kin-directed punishment is stronger in intensive kinship systems. The exploratory results further suggest that kin-directed punishment declines as societies develop greater technological and social differentiation, with this association operating both directly (TSD's robust direct effect) and indirectly through the RI → TSD pathway. Community size shows no consistent association.

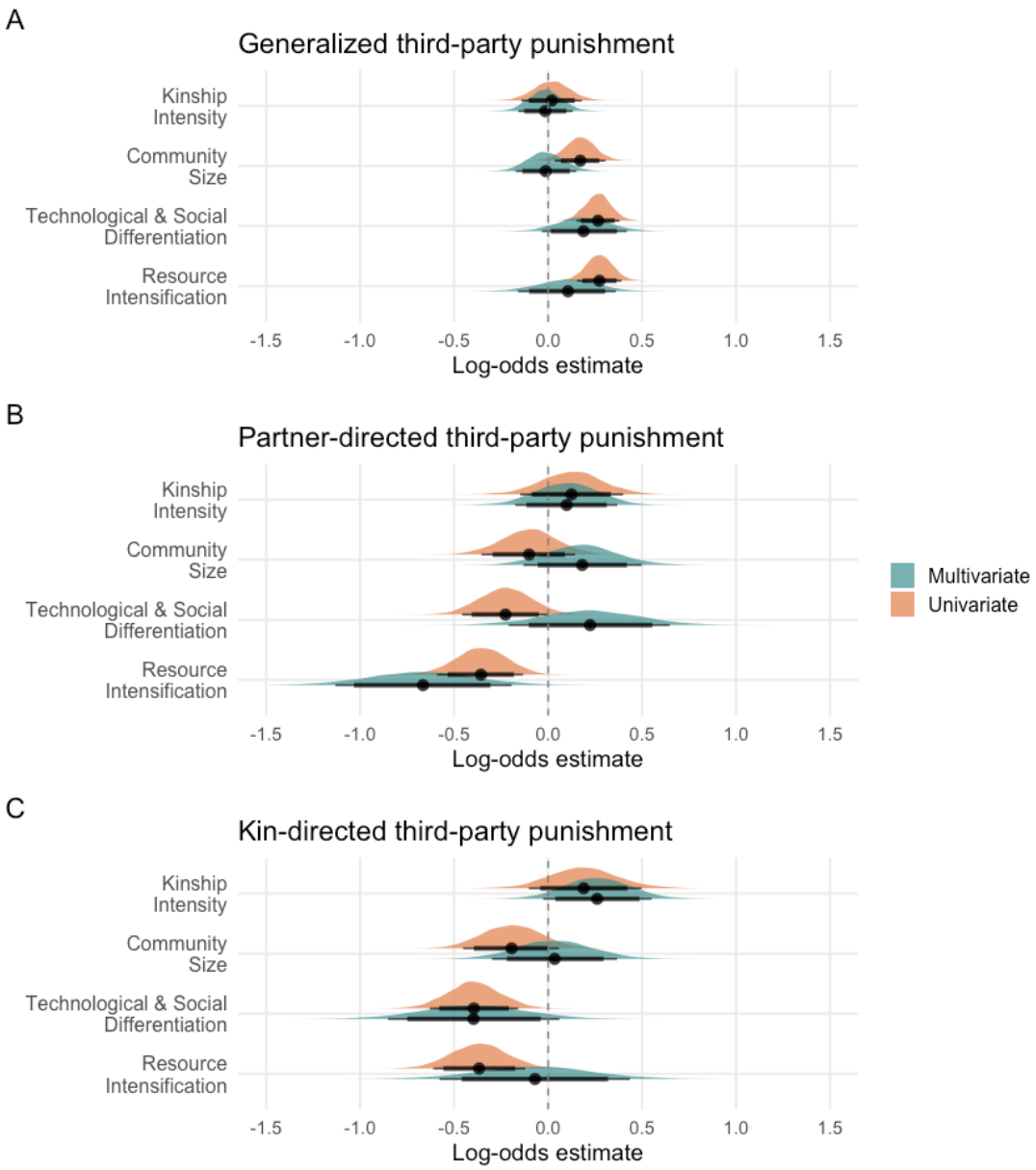

Figure 3: Predictors of third-party punishment subtypes.

Table 1: Preregistered hypothesis scorecard. Predicted: direction predicted in preregistration (https://osf.io/79g4e). Observed: direction of multivariate posterior median. Support: classification based on multivariate model. 'Supported' = 90% HDI excludes 0 in predicted direction; 'Moderately supported' = 80% HDI excludes 0 in predicted direction; 'Suggestive (NS)' = posterior median in predicted direction with pd > 0.80 but 80% HDI includes 0; 'Not supported' = posterior median against predicted direction or pd

Outcome	Predictor	Predicted	Observed	β [90% HDI]	Support
2PP	RI	↓	↑	0.04 [-0.23, 0.33]	Not supported
	TSD	↓	↓	-0.32 [-0.57, -0.07]	Supported
	CS	↓	↓	-0.02 [-0.20, 0.16]	Not supported

Table 1: Preregistered hypothesis scorecard. Predicted: direction predicted in preregistration (https://osf.io/79g4e). Observed: direction of multivariate posterior median. Support: classification based on multivariate model. 'Supported' = 90% HDI excludes 0 in predicted direction; 'Moderately supported' = 80% HDI excludes 0 in predicted direction; 'Suggestive (NS)' = posterior median in predicted direction with pd > 0.80 but 80% HDI includes 0; 'Not supported' = posterior median against predicted direction or pd

Outcome	Predictor	Predicted	Observed	β [90% HDI]	Support
	KII	↑	↓	-0.02 [-0.18, 0.15]	Not supported
3PP (aggregate)	RI	↑	↓	-0.03 [-0.27, 0.20]	Not supported
	TSD	↑	↑	0.13 [-0.06, 0.35]	Suggestive (NS)
	CS	↑	↑	0.01 [-0.14, 0.15]	Not supported
	KII	↓	↑	0.04 [-0.10, 0.17]	Not supported
3PP-General	RI	↑	↑	0.11 [-0.16, 0.36]	Not supported
	TSD	↑	↑	0.18 [-0.03, 0.42]	Moderately supported
	CS	↑	↓	-0.02 [-0.17, 0.15]	Not supported
3PP-Partner	RI	↑	↓	-0.67 [-1.14, -0.19]	Not supported
	TSD	↑	↑	0.22 [-0.21, 0.65]	Suggestive (NS)
	CS	↑	↑	0.18 [-0.13, 0.50]	Suggestive (NS)
3PP-Kin	KII	↑	↑	0.26 [-0.03, 0.55]	Moderately supported

Exploring socio-behavioral contexts of punishment

To investigate the cultural and behavioral domains associated with different punishment types, we conducted exploratory analyses using Outline of Cultural Materials (OCM) subject codes[64] as predictors in Bayesian multilevel models with regularized horseshoe priors, which allow systematic identification of associated codes while shrinking irrelevant predictors toward zero. See Figures 4 and 5 and Supplementary Table 4 for all model coefficients.

For 2PP, the OCM codes with the strongest positive associations included *Family Relationships*, *Sex and Marital Offenses*, *Ingroup Antagonisms*, and *Extramarital Sex Relations*. Negative associations were observed for *Community Councils*, *Kinship Regulation of Sex*, *Purification and Atonement*, and *External Relations*.

For 3PP, the OCM codes with the strongest positive associations included *Special Unions and Marriages*, *Celibacy*, and *Taxation and Public Income* (which fall outside the original

search-seed code set), as well as *Social Offenses*, *Sanctions*, *Offenses Against the State*, *Police*, and *Execution of Justice* (which were among the search-seed codes).

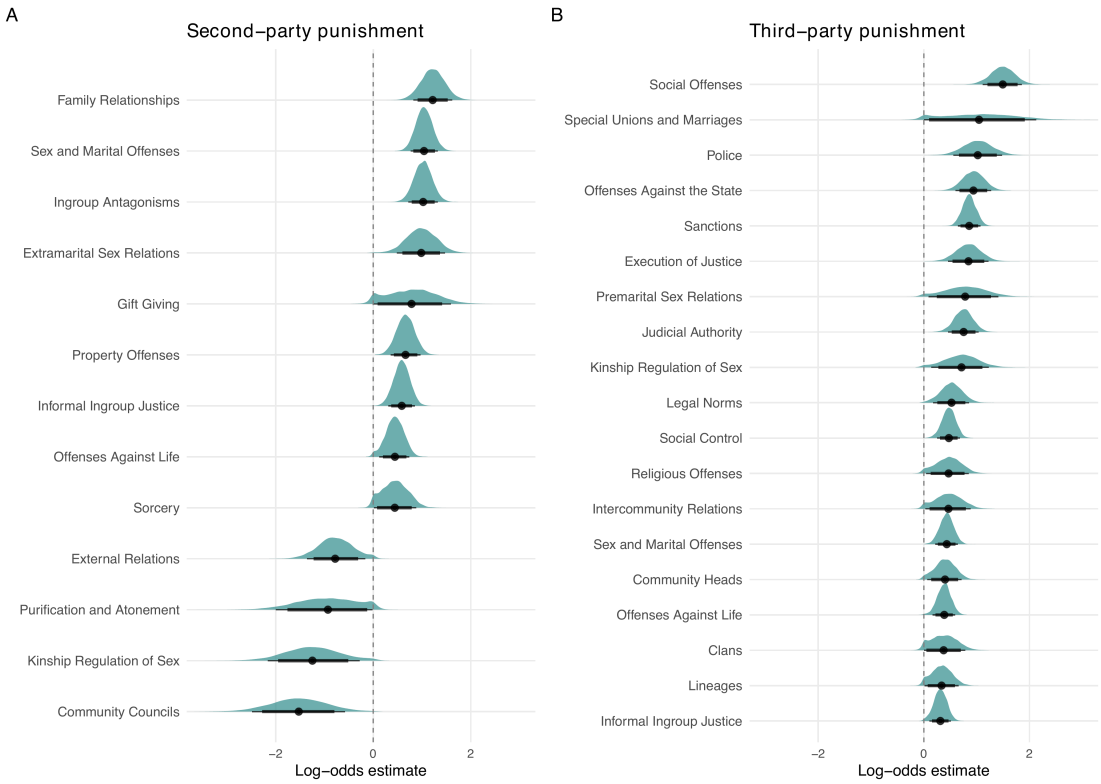

Figure 4: OCM code effects on second- and third-party punishment. Points show posterior medians with 80% and 90% credible intervals. Only codes with 90% credible intervals excluding zero are shown.

Among the three subtypes of third-party punishment, interesting patterns emerged. 3PP-Kin was most strongly associated with *Kin Relationships*, *Offenses Against Life*, *Informal Ingroup Justice*, *Sorcery*, *Lineages*, and *Clans*. Negative associations were observed for *Wrongs*, *Property Offenses*, and *Judicial Authority*.

3PP-Partner showed positive associations with *Friendships*, *Clans*, and *Offenses Against Life*.

3PP-General was most strongly associated with *Social Offenses*, *Police*, *Religious Denominations*, *Execution of Justice*, and *Offenses Against the State*. Other positive predictors indicate associations with legal norms and hierarchical and institutionalized leadership structures. Negative associations were observed for *Kin Relations*, *Mourning*, and *Family Relations*.

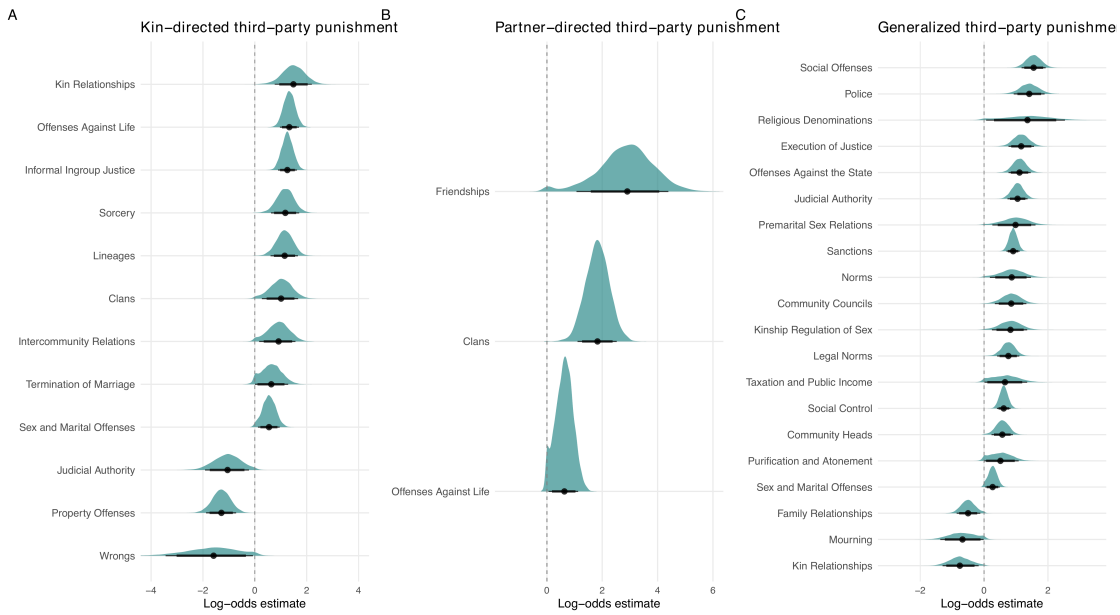

Figure 5: OCM code effects on third-party punishment subtypes. Points show posterior medians with 80% and 90% credible intervals. Only codes with 90% credible intervals excluding zero are shown.

Discussion

Using ethnography-derived data from 60 culturally diverse societies, we find that third-party punishment is both more widespread and more varied than existing accounts suggest. Challenging claims that norm-enforcing, generalized third-party punishment is rare or weakly developed in smaller-scale societies, our results indicate that such punishment occurs across a wide range of human societies, including among hunter-gatherers and in the absence of formal institutions. At the same time, its social form systematically varies: generalized third-party punishment is associated with greater technological and social differentiation and, to a lesser extent, resource-use intensification, whereas kin-directed punishment is linked to kinship intensity and lineage-based organization. Second-party punishment, by contrast, tends to decline with increasing differentiation. These findings challenge conceptions of punishment as a unitary phenomenon and suggest that understanding its role in human cooperation requires distinguishing among relational forms and situating them within their broader socioecological contexts. Below, we elaborate on each of these findings in turn.

Third-party punishment without institutional enforcement

Ethnographic research has often emphasized that generalized third-party punishment, particularly in the form of corporal punishment, execution, or material sanctions, is uncommon in smaller-scale societies[13,16,27,28,60,65]. This view has shifted explanatory weight toward dyadic enforcement, including retaliation, demand-sharing, and bargaining driven by prosocial preferences or self-interested concerns for equality[11,13,66].

Our findings complicate this framework. Generalized third-party punishment appears repeatedly in ethnographic descriptions across our sample, including among hunter-gatherers and outside formal institutional settings. Preliminary coding from an ongoing follow-up study using the paragraph-level dataset indicates that, among 1,176 texts coded as containing evidence for generalized third-party punishment, roughly 38% reflect institutionalized enforcement, 27% non-institutionalized enforcement, and the remainder are ambiguous or reflect both. Non-institutionalized third-party punishment thus appears in the ethnographic record at a frequency comparable to institutionalized third-party punishment.

These descriptive patterns align with ethnographic accounts in which third parties intervene in norm enforcement in the absence of formal courts or codified legal systems. In De Coccola's account of the Copper Inuit in the North American arctic, homicide triggers a socially organized enforcement process in which "an avenger has been chosen" and the offender becomes socially marked for execution[67]. Among the Tlingit in southern Alaska, incest violations are sanctioned through clan authority, with offenders killed by their respective clansmen[68]. Turnbull's description of Mbuti social life in the Congo Basin of Central Africa emphasizes camp-level punishment carried out through ridicule, ostracism, and occasionally physical sanction driven by "the weight of public opinion"[69]. Similarly, Gusinde describes execution of dangerous offenders as communal amends rather than private revenge among the Ona of Tierra del Fuego in South America[70]. These cases illustrate that third-party punishment occurs without formal institutions, emerging instead through collective authorization and shared moral evaluation.

Across the full range of societies in our sample, however, we do find evidence that third-party punishment is positively associated with formal authority and institutionalized sanctioning. Our exploratory OCM-domain analysis (Fig. 4B) reveals positive associations with domains such as *Social Offenses*, *Sanctions*, *Offenses Against the State*, *Legal Norms*, *Execution of Justice*, and *Police*. This pattern aligns with findings from a similar ethnographic study which systematically analyzed texts on leadership across 59 societies and found that punishment was described as a function of leaders in approximately half of the sample. Notably, this association was not explained by socioecological factors such as subsistence type or group context[24]. Evidence for leader-enacted punishment was also the strongest predictor of coercive leadership across a suite of 109 leadership dimensions[24].

These results indicate that third-party punishment is neither confined to formal judicial institutions nor absent from smaller-scale societies. Rather, it spans a continuum. In some contexts it emerges through decentralized, kin- or community-based enforcement. In others it is embedded within more formalized systems of legal authority and either coercive or collective governance. Operationalizations that equate third-party punishment solely with institutional courts risk overlooking socially coordinated yet noninstitutionalized enforcement[71,72].

Kinship intensity and kin-directed third-party punishment

This shift in focus highlights a second question: how socioecological structure shapes the relational targets of third-party punishment. Our analyses reveal a modest positive association between kinship intensity and third-party punishment enacted on behalf of kin (Fig. 3C). Ethnographic evidence suggests that in these contexts, punishment is frequently carried out through corporate kin groups, lineage representatives, or delegated relatives.

Among the Tlingit, homicide disputes are framed as conflicts between clans, and punishment may involve executing a member of the offender's clan to prevent escalation[73]. Among Tiv horticulturalists in Nigeria homicide cases similarly involve collective deliberation, with punishment delegated to senior kin representatives[74]. Among the Saramaka Maroons of Suriname, a husband whose wife has committed adultery may seek redress through his wife's lineage, which may "either authorize the husband and his kinsmen to give the man a beating or help to extract a payment from him"[75]. Kin groups vary in their enforcement policy. Some favor beatings, others payments, and some prefer no action, illustrating how kin-based third-party punishment is not automatic but mediated through lineage-level deliberation and collective authorization[75]. Among Wolof communities in Senegambia, relatives of individuals believed to have been harmed through witchcraft may coerce the suspected offender to repair the harm, sometimes forcing him to "restore the part of the body he had taken, and perhaps give medicines to aid in the recovery"[76]. Among Bedouin pastoralists in North Africa, accounts emphasize lineage-based retaliation carried out by delegated kin actors[77,78]. In northern Syria and Turkey Kurdish honor-based sanctions further illustrate punishment enacted by relatives acting as guardians of collective reputation[79].

Our exploratory OCM-domain analysis supports this interpretation. Kin-directed third-party punishment shows positive associations with domains related to *Offenses Against Life*, *Kin Relationships*, *Clans*, and *Lineages*, while showing comparatively weaker or negative associations with domains emphasizing *Judicial Authority* or generalized "wrongs" (Fig. 5A). These patterns reinforce the view that kin-directed 3PP represents a form of lineage-based enforcement–socially regulated but operating outside institutionally centralized judgement.

These cases clarify what kinship intensity captures in our models. In kin-intensive societies, punishment frequently operates through corporate kin responsibility and delegated enforcement rather than dyadic retaliation or impersonal institutions. This modest but consistent pattern suggests that intensive kinship structures facilitate third-party punishment by establishing corporate kin groups with recognized authority to adjudicate offenses, enforce collective liability, and intervene on behalf of members–a structurally distinct mode of third-party enforcement that operates through kinship institutions rather than either dyadic retaliation or impersonal authority. Notably, technological and social differentiation showed a consistent negative association with kin-directed punishment in both univariate and multivariate models, with the 90% credible interval narrowly crossing zero. This pattern suggests that as societies develop more

specialized institutional structures, enforcement shifts away from lineage-based mechanisms. This pattern mirrors the positive association between differentiation and generalized third-party punishment discussed below.

Societal differentiation and generalized third-party punishment

In contrast, generalized third-party punishment is associated with technological and social differentiation (Fig. 3A). Although resource-use intensification was also associated with generalized 3PP in univariate analyses, this association attenuated substantially when TSD was included, consistent with RI's effect on generalized 3PP operating largely through TSD-mediated pathways rather than through a direct effect on enforcement. This interpretation aligns with Ringen et al.'s[61] finding that resource-use intensification causally drives technological and social differentiation over evolutionary time, but not vice versa: intensified production systems appear to facilitate generalized 3PP primarily by promoting the specialized roles, formal procedures, and political integration that constitute TSD, which in turn support institutionalized norm enforcement.

Ethnographic accounts from high-TSD societies describe punishment being carried out through specialized roles, formal procedures, and routinized sanctions. Leach's account of the Kachin agriculturalists of highland Myanmar describes courts that impose fines or expulsion following formal adjudication[80]. Hulbert's description of early twentieth-century Korean agricultural communities documents penal codes specifying how beatings are administered and regulated by law[81]. DeGlopper's ethnography of Taiwanese maritime trading communities describes judicial procedures involving detention, graded sentencing, confiscation, and corporal punishment[82]. Other examples include bureaucratic disciplinary systems among state-linked agricultural populations in central Thailand[83], judicial punishment administered by chiefs and elders in Akan communities of West Africa[84], and religious authorities empowered to enforce sanctions in Afro-Brazilian urban communities[85]. In these contexts, punishment is a socially differentiated function enacted through recognized offices or institutions rather than kin-based or interpersonal mechanisms.

These cases help explain why generalized third-party punishment covaries with societal differentiation in our models. Rather than indicating more punishment overall, the association reflects the emergence of procedurally formalized and role-specialized enforcement systems[86].

Resource-use intensification and partner-directed third-party punishment

Our results for partner-directed third-party punishment did not match our preregistered prediction. We expected partner-based 3PP to increase with sociocultural complexity; particularly with technological and social differentiation and community size, where individuals depend on durable, valuable, and difficult-to-replace partnerships. Instead, we found a robust negative association between resource-use intensification and partner-directed 3PP, an association that strengthened in the multivariate specification (Fig. 3B). Following the framework outlined in the Methods, the strengthening of RI's negative

coefficient when TSD is included suggests that RI exerts a direct negative effect on partner-based enforcement, partially offset in the univariate specification by a small positive contribution running through TSD.

Several substantive interpretations of this unexpected pattern are possible. One possibility is that resource-use intensification, by tying individuals to fixed productive resources (intensive agricultural plots, stored food surpluses, defensible territories), reduces the relational mobility that makes durable partnerships strategically valuable. In foraging and horticultural societies, alliances with non-kin partners may carry high stakes because individuals can leave one social context and seek protection in another; in intensified production systems, exit options are constrained by attachment to land and stored resources, and partner-based enforcement may give way to more institutionalized or kin-based mechanisms.

A second possibility concerns operationalization. Our partner-based 3PP category requires that the punisher and victim share a recognized non-kin alliance going beyond co-residence. This category may be harder to identify ethnographically in densely interconnected societies where most relationships are simultaneously kin, neighbors, and ritual partners, leading partner-directed enforcement to be coded under other categories. Future work using more fine-grained operationalizations of partnership relations, particularly in ethnographic contexts where such relationships are clearly demarcated (e.g., trading partnerships, ritual exchange relationships, formalized friendships), could help clarify whether the pattern we observe reflects a genuine substantive finding or a measurement limitation.

We flag this as an unexpected and exploratory finding requiring further investigation rather than as a definitive result. The robustness of the negative RI association across specifications suggests it is unlikely to be statistical noise, but its substantive interpretation should be tested with additional cross-cultural and within-society evidence.

A context-dependent account of punishment

These findings help reconcile competing theoretical accounts of third-party punishment. Strong reciprocity theory predicts that individuals will unilaterally enforce social norms even at personal cost, driven by internalized punishment preferences[9,34]. Coalitional control accounts emphasize instead that norm enforcement is collective and socially coordinated, with groups acting together to suppress dominance and regulate deviance[87]. A third tradition questions both, highlighting ethnographic evidence that punitive third-party enforcement is rare in small-scale societies and is often replaced by mediation, compensation, or dyadic resolution[13,28,65]. Our results speak differently to each. They sit uneasily with strong reciprocity, since the generalized third-party punishment we document is predominantly collective and communally authorized rather than unilaterally initiated by individual moralistic punishers. They are more consistent with coalitional control in that norm enforcement appears to operate through shared social authorization. And they complicate the view that third-party punishment is rare: generalized enforcement

is widespread in our sample, but its prevalence varies systematically with socioecological conditions. Studies focused on smaller-scale, less differentiated societies are more likely to encounter kin-directed or partner-based enforcement than generalized punishment, because the latter likely emerges with technological and social differentiation. The apparent rarity of punitive third-party enforcement in such societies thus reflects socioecological variation rather than an absence of third-party punishment per se–a pattern obscured when third-party punishment is treated as a unitary phenomenon.

Models of strong reciprocity emphasize individuals' willingness to incur personal costs to punish norm violations, even absent future material or reputational rewards[9,34]. Our data cannot directly speak to punishers' motivations or whether enforcement is materially costly in the contexts we document. Indeed, many ethnographic cases of third-party punishment involve reputational consequences, reciprocal obligations, or institutional roles that likely offset or obscure individual costs. What our findings do contribute is evidence that third-party punishment is not rare in the ethnographic record, as some accounts have suggested, but that its expression is patterned by social structure rather than uniform across societies. This is relevant for experimental and theoretical work. Laboratory paradigms that model punishment as a context-free individual decision may miss the relational and institutional channels through which punitive behavior is organized in natural settings[13,55].

Coercive enforcement accounts offer a complementary lens, emphasizing that punishment systems reflect the distribution of power and authority within societies rather than prosocial motivations alone[20,22,87,88]. Our results align with this perspective. Where authority is concentrated in corporate kin groups, punishment tends to be organized through lineage-based responsibility; where enforcement capacity is distributed across specialized roles and institutions, punishment becomes more generalized and procedurally mediated. These patterns suggest that understanding cross-societal variation in punishment requires attending to who holds enforcement authority and how that authority is socially organized. These questions link the evolutionary study of cooperation to institutional analysis in economics and political science[89,90] and to relational models of norm enforcement in psychology[22,32,91,92], and suggest that cross-disciplinary synthesis may be needed to fully account for the diversity of punishment systems documented here.

Rather than a single continuum from simple to complex societies, the ethnographic record points to multiple punishment regimes. Kin-intensive systems support kin-directed punishment grounded in lineage obligations, whereas social differentiation facilitates generalized punishment enacted through institutional roles and leadership structures.

Limitations and future directions

This study has several limitations. Our operationalization of punishment is intentionally broad, including physical sanctions, social exclusion, ridicule, delegated violence, and institutionally administered penalties. Future research could differentiate sanction types (e.g., corporal punishment, gossip, ostracism, economic sanctions, ritual humiliation) to

assess whether they exhibit distinct socioecological patterning[20]. Likewise, although we distinguish relational targets of third-party punishment, we do not fully disaggregate enforcement agents, such as uninvolved community members, delegated kin representatives, recognized leaders, councils, or formal judicial institutions. Separating these actor categories would allow finer-grained analysis of how authority structures shape enforcement. More generally, bridging ethnographic evidence with behavioral, economic, and moral-psychological approaches will require modeling punishment along multiple dimensions: relational target, enforcement agent, institutional embedding, and cost structure. Doing so will better align evolutionary and behavioral theories of cooperation with the diversity of enforcement systems documented across human societies.

Our study also faces limitations inherent in ethnography-based analyses[93–95]. Ethnographic texts are not perfect records of social life but are shaped by ethnographers' theoretical commitments, research interests, and subjectivity. The ethnographic record is also likely systematically biased toward observable, public behaviors (overt punishment in our case), which may be reported more frequently than subtle or private forms of sanctioning. Unmeasured features of ethnography or ethnographers, such as theoretical orientation, rapport with communities, or duration of fieldwork, may also influence available evidence. The ethnographic record is further subject to the "absence of evidence" problem, where the lack of mention of particular punishment practices does not necessarily indicate their absence in a society. We therefore frame our conclusions as descriptive patterns within the ethnographic record rather than definitive claims about universal features of human punishment. Results reported here should be interpreted in conjunction with long-term and individual-level data from naturalistic communities, experimental evidence, and formal models.

Punishment is not a single solution to the challenges of cooperation, but a diverse repertoire of norm enforcement practices shaped by the socioecological contexts in which human groups are situated[22,96]. By moving beyond simple distinctions between second- and third-party punishment to examine who punishment defends (e.g., kin, partners, or community members) we reveal how kinship intensity and societal and economic differentiation differentially pattern enforcement systems. These findings underscore the value of multidimensional approaches to punishment that attend to relational targets, enforcement agents, and institutional context. Building a more complete account of how cooperation is sustained across human societies will ultimately require integrating such ethnographic evidence with experimental, behavioral, and formal-modeling approaches.

Methods

Ethnographic sample and coding

Our data collection process involved extracting a sample of 15380 paragraphs from the electronic Human Relations Area Files World Cultures database (eHRAF), a digitized repository of ethnographic documents indexed at the paragraph level using the Outline of

Cultural Materials (OCM) subject classification system (https://ehrafworldcultures.yale.edu). We restricted our analysis to documents from the 60-culture Probability Sample Files (PSF), which was designed to mitigate Galton's problem of cultural non-independence by randomly sampling one society with high-quality ethnographic coverage from each of 60 world regions[63]. Our culture sample is given in Figure 6. To identify text passages potentially describing offenses and responses to offenses, we conducted an advanced search using 13 OCM codes most closely linked to punishment, offense, and norm violation: *Social Control* (626), *Informal In-Group Justice* (627), *Execution of Justice* (696), and all codes within the *Offenses & Sanctions* major category (681–689), which covers sanctions, offenses against life, offenses against the person, sex and marital offenses, property offenses, nonfulfillment of obligations, offenses against the state, religious offenses, and social offenses. We used the Boolean operator OR across all 13 codes, returning any paragraph from any document within the 60 PSF societies associated with at least one of these codes. This search returned 15380 paragraphs (at the time of data collection, circa 2021), which formed the basis of our subsequent coding procedures.

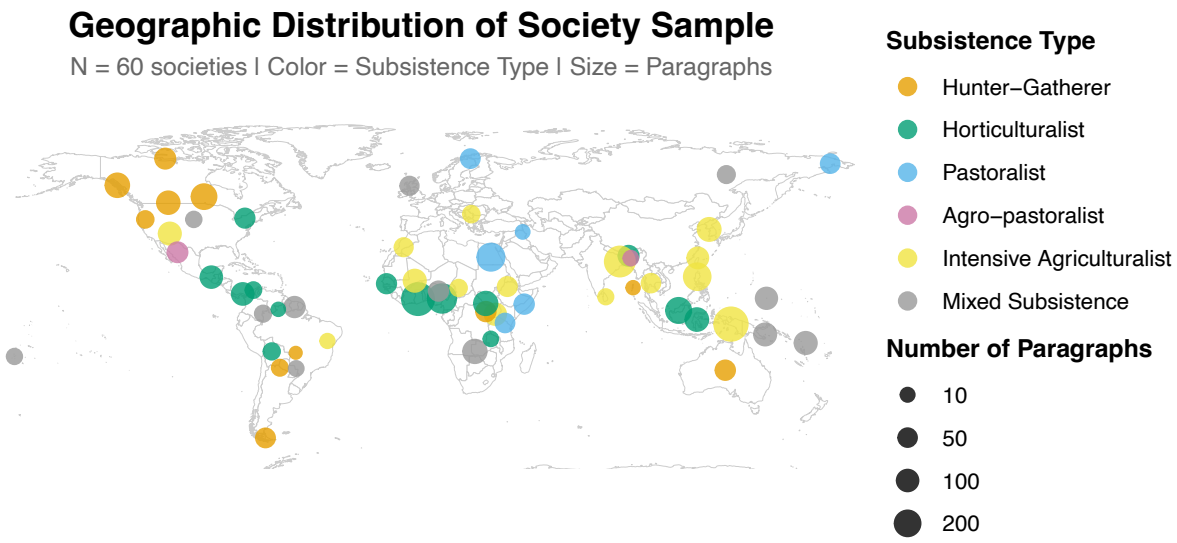

Figure 6: Geographic distribution of the 60-society probability sample. Point size reflects the number of coded paragraphs per society; color indicates subsistence type.

We then developed a multi-stage coding process to produce quantitative measures of our variables. First, we used a random sub-sample of 150 paragraphs to refine the operational definitions of the coding scheme and develop a master training dataset coded and agreed upon by the project PIs (ZHG, CM, FDP, and former PI AM). Second, we recruited research assistants, each of whom completed a pre-coding training session, independently coded the 150 training paragraphs, and then completed a post-coding calibration session comparing their codes to the master training dataset to identify and correct systematic differences (inter-coder reliability statistics for each coder are reported in the Supplementary Materials). Third, we assigned coders random subsets of the full paragraph sample for independent coding. This process yielded 3961 paragraphs coded by at least two coders, of which 1071 were coded by three coders. Across these paragraphs and 6 coded variables, coders produced 11893 individual paragraph-level codings.

We defined punishment broadly as any action that imposes a cost on another party on the basis of an offense or violation of a social norm. Costs could include physical harm (e.g., beating, execution), material sanctions (e.g., fines, confiscation), social sanctions (e.g., ostracism, exclusion), verbal sanctions (e.g., ridicule, public condemnation), or other forms of socially recognized harm. Second-party punishment (2PP) was coded when the victim of an offense acted to sanction the offender. Third-party punishment (3PP) was coded when an individual or group other than the victim imposed a cost on the offender. Where punishment was clearly present but insufficient information was available to determine whether it was second- or third-party, it was coded as unidentified punishment (UPP). When third-party punishment was identified, we further classified it according to whether it was enacted on behalf of kin, social partners, or the broader community.

All inferential and descriptive statistical analyses were restricted to paragraphs independently coded by at least two coders to ensure reliability of paragraph-level classifications. To account for the nested structure of the data (i.e., multiple coder evaluations per paragraph, paragraphs nested within documents, and documents nested within societies) we used Bayesian multilevel models throughout, as described below.

Table 3: Researcher-coded variables and operational definitions.

Variable	Operational definition
Second-party punishment (2PP)	Across all offenses mentioned in the text record, is there evidence for second-party punishment associated with any offense? Second-party punishment is when the victim of the offense acts to punish the offender.
Third-party punishment (3PP)	Across all offenses mentioned in the text record, is there evidence for third-party punishment associated with any offense? Third-party punishment is when someone other than the victim of the offense acts to punish the offender.
Third-party punishment on behalf of kin (3PP: kin)	Did the third-party punisher punish the offender on behalf of their kin, i.e., are the punisher and victim kin or family members?

Table 3: Researcher-coded variables and operational definitions.

Variable	Operational definition
Third-party punishment on behalf of partners (3PP: partner)	Did the third-party punisher punish the offender on behalf of a social partner, i.e., are the punisher and victim friends, allies, or have some positive social relationship, beyond simple co-residence relationships?
Third-party punishment on behalf of general community members (3PP: general)	Did the third-party punisher punish the offender on behalf of a general community member, i.e., are the punisher and victim unrelated and/or unacquainted, but also general community members?
Unidentified punishment (UPP)	Across all offenses mentioned in the text record, is there evidence of punishment associated with any offense, but which cannot be defined as second-party or third-party punishment? There is evidence of punishment occurring, but not enough information to determine if it is second or third-party.

Statistical analyses

All analyses were conducted in R using Bayesian models fit with brms[97] and the cmdstanr backend[98]. The primary unit of analysis was the coder-by-paragraph coding. We restricted the analytic dataset to ethnographic paragraphs that were independently coded by at least two coders, and we analyzed each punishment category as a binary outcome indicating whether a coder recorded "evidence for" the focal category (1) versus any other code, i.e., "no evidence" (0). Outcomes included second-party punishment (2PP), third-party punishment (3PP), third-party punishment on behalf of kin (3PP-kin), third-party punishment on behalf of social partners (3PP-partner), generalized third-party punishment (3PP-general), and unspecified punishment.

Coding-level support

To estimate how frequently the ethnographic record contained evidence for each punishment type at the coding level, we fit intercept-only Bayesian multilevel logistic regression models with a Bernoulli likelihood and logit link. Each model included varying intercepts for paragraph (`uuid`), coder (`coder`), document (`document_id`), and society (`owc_id`). This structure accounts for repeated ratings of the same paragraph, coder-specific response tendencies, clustering of paragraphs within documents, and heterogeneity across societies. We summarized posterior distributions of the population intercept on the probability scale via the inverse-logit transform, yielding posterior estimates of the proportion of coder-by-paragraph ratings supporting each punishment category.

Culture-level support

Because coding-level prevalence can be sensitive to unequal documentation across societies, we also estimated culture-level prevalence of each punishment category. We first aggregated coder ratings within paragraph by counting the number of coders marking "evidence for" each category. We treated a paragraph as providing "confirmed" evidence for a category when at least two coders independently recorded "evidence for." We then collapsed to the society level by defining a binary indicator (present) equal to 1 if a society contained at least two paragraphs with confirmed evidence for the category and 0 otherwise. We selected the two-paragraph threshold to require evidence beyond a single isolated case while remaining permissive enough to detect punishment types that may be reported infrequently in the ethnographic record. Culture-level prevalence was estimated with intercept-only Bernoulli models and summarized on the probability scale.

Hypothesis tests with culture-level predictors

We tested preregistered hypotheses (https://osf.io/79g4e) linking socioecological structure to punishment categories using Bayesian multilevel logistic regression models predicting coder-by-paragraph evidence from society-level predictors. Predictors comprised kinship intensity (KII) calculated using methods based on Schulz et al.[62] and three measures of societal differentiation and scale: resource-use intensification (RI), technological and social differentiation (TSD), and community size/urbanization (CS; Urbanization_SCCSv152). RI and TSD were treated as latent dimensions following the measurement model validated by Ringen et al.[61], in which RI captures variation in agricultural intensification, fixity of residence, population density, urbanization, reduced reliance on hunting, and food storage, while TSD captures variation in writing, labor specialization, money, political integration, and social stratification (Fig. 7). We re-estimated these latent dimensions on a pooled sample comprising the full Standard Cross-Cultural Sample (SCCS)[99] plus the PSF societies not included in the SCCS, for which we hand-coded the necessary indicator variables, preserving the original indicator structure to maintain comparability with the Ringen et al. measurement model. Our updates to the Ringen et al. code are available on GitHub (https://github.com/zhgarfield/complex_coev_sccs). Predictor values were merged to coded paragraphs via society identifier (`owc_id`).

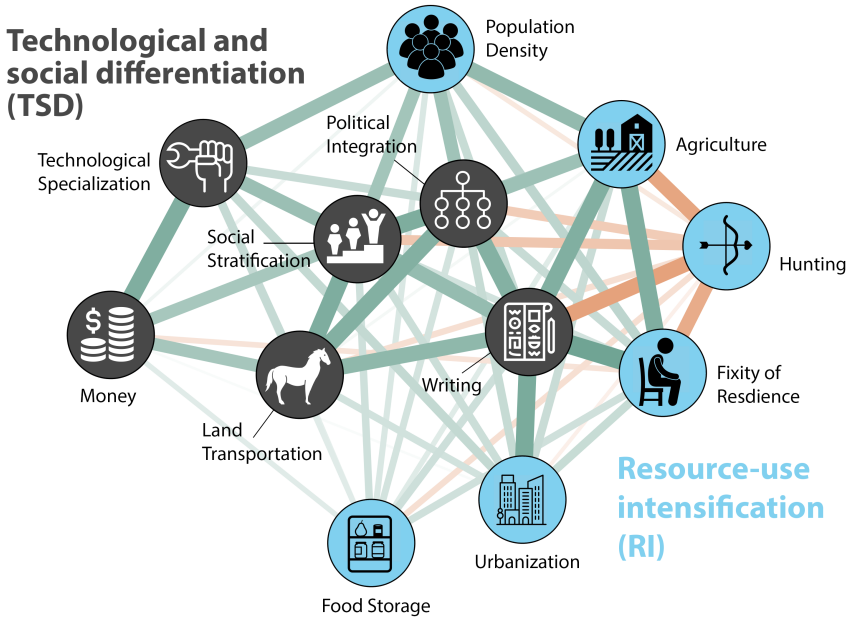

Figure 7: Measurement model for resource-use intensification (RI) and technological and social differentiation (TSD) as validated by Ringen et al. (2021). RI loads onto agriculture, fixity of residence, population density, urbanization, hunting (negative), and food storage. TSD loads onto writing, labor specialization, money, political integration, and social stratification. Figure reprinted from Ringen et al. (2021).

RI and TSD are conceptually distinct dimensions whose discriminant validity was established by Ringen et al.[61], who demonstrated that a two-factor model substantially outperforms a unidimensional "complexity" model in predicting the global distribution of nine indicators across 186 SCCS societies (ΔELPD = 75.4 favoring the two-factor model). Although the two factors are correlated–Ringen et al. report a phylogenetic correlation of ρ = 0.64 [90% HPDI: 0.41, 0.84]–they capture conceptually and empirically separable aspects of socioecological organization. Ringen et al. further established a causal ordering between them using a dynamic phylogenetic coevolutionary model, finding that increases in RI lead to increases in TSD over evolutionary time but not vice versa. We retain RI and TSD as separate predictors in our models to preserve this established theoretical and empirical distinction, and we interpret coefficients accordingly.

In our PSF subsample, RI and TSD are correlated at r = 0.88. Given Ringen et al.'s evidence that RI is causally upstream of TSD, the univariate and multivariate specifications we report estimate distinct quantities. Univariate coefficients on RI estimate its total association with each punishment outcome, encompassing both direct effects and effects mediated through the RI → TSD pathway. Multivariate coefficients estimate RI's direct association after partialing out the TSD-mediated component. Where univariate and multivariate estimates agree, the finding is robust to whichever specification one prefers; where they diverge, the divergence is informative about the relative contribution of direct versus mediated pathways. We therefore report both specifications throughout.

For each punishment outcome, we fit a multivariate model including all four standardized predictors simultaneously, with varying intercepts for paragraph (`uuid`), coder (`coder_id`), document (`document_id`), and society (`owc_id`). Predictors were z-scored within each outcome-specific dataset prior to model fitting; fixed-effect coefficients are therefore interpretable as the change in the log-odds of "evidence for" associated with a one standard deviation increase in the predictor. To assess robustness to predictor intercorrelations, we additionally fit univariate models in which each predictor was entered separately, retaining the same multilevel structure. Models were estimated using Hamiltonian Monte Carlo as implemented in brms with the cmdstanr backend (4 chains, 4,000 iterations per chain, adapt_delta = 0.95) and weakly regularizing priors: $\mathcal{N}(0,1)$ for fixed effects and $Exponential(1)$ for group-level standard deviations.

To verify robustness to deep cultural-historical relatedness among societies, we additionally refit each multivariate model with a varying intercept for Glottolog language family[100]. Language family was assigned via mapping each PSF society's HRAF identifier to its Glottolog language code, then to the language's top-level family. Of the 60 societies, 33 distinct families were represented, including 24 singletons; the largest families were Atlantic-Congo (n = 8), Austronesian (n = 7), and Indo-European (n = 6). Robustness model coefficients are reported alongside primary model coefficients in Supplementary Table 8.

Exploratory analyses using the eHRAF-provided data

We leveraged the Outline of Cultural Materials (OCM) codes provided by eHRAF as predictor variables to identify the topics and features commonly associated with ethnographic descriptions of second- and third-party punishment. The OCM coding scheme is an ethnographic classification system of human behavior, social life and customs, material culture, biology, and human-ecological environments, designed to capture a wide range of topics relevant to the human sciences. The core feature of the eHRAF World Cultures database is the association of each paragraph with relevant OCM codes discussed in the text. The OCM scheme is hierarchically organized and includes more than 90 major categories (such as 550 *Social Stratification* and 740 *Health and Welfare*) which contain more than 700 individual codes (such as 554 *Status, role, & prestige*; 744 *Public Health & Sanitation*). For each paragraph in our sample, we extracted all associated OCM codes, yielding a binary matrix indicating the presence or absence of each code. After removing bibliographic, geographic, and methodological codes (e.g., *Literary texts*, *Texts translated into English*, *Ethnographic data*, *Coding notes*, *location*) and codes appearing in fewer than three paragraphs, our final predictor set included 345 individual OCM codes.

Our final predictor set retains the 13 OCM codes used in the initial Boolean retrieval (i.e., *Social Control*, *Informal In-Group Justice*, *Execution of Justice*, and the *Offenses & Sanctions* major category, codes 681–689). These codes are necessarily prevalent in our analytic sample because they were used to construct it, and on a single-outcome analysis their presence as predictors would be circular. We retain them across our five outcome models because they vary differentially by punishment subtype: a paragraph coded *Police*

or *Execution of Justice* may signal generalized 3PP but not kin-directed or partner-directed enforcement, and the relative discrimination of search-seed codes across subtypes is itself an informative pattern. Coefficients for these codes should be interpreted as describing how broad punishment-relevant cultural domains differentially track distinct punishment subtypes, not as evidence that "punishment-related codes predict punishment." Coefficients for codes outside the search-seed set (e.g., *Friendships*, *Clans*, *Lineages*, *Sorcery*, *Mourning*, *Family Relationships*) carry stronger inferential weight as independent predictors.

Given the relatively large number of OCM codes available as predictor variables (345 codes) and the paragraph-level structure of our data, we employed Bayesian multilevel models with regularized horseshoe priors to identify OCM codes predictive of evidence for punishment at the paragraph level. The regularized horseshoe prior is a continuous shrinkage prior particularly useful when the number of predictors is large relative to the number of cases, as it helps address multicollinearity and reduces the potential for overfitting by shrinking irrelevant coefficients toward zero while allowing relevant coefficients to escape shrinkage[101,102]. This approach identifies OCM codes that are systematically associated with evidence for each punishment type while accounting for the nested structure of our data through random intercepts for paragraph and coder.

All OCM code predictors were centered (but not scaled) prior to model fitting. We used weakly regularizing priors: $\mathcal{N}(0,1)$ for the intercept, $Exponential(1)$ for random effect standard deviations, and a regularized horseshoe prior (df = 1, par ratio = 0.1) for the OCM code coefficients. The par ratio parameter controls the expected sparsity, with smaller values inducing stronger shrinkage; our choice reflects the expectation that only a subset of OCM codes would be meaningfully associated with each punishment outcome.

Models were fit using Hamiltonian Monte Carlo with four chains, each running 2,000 iterations (including 1,000 warmup). We assessed convergence using R-hat statistics (< 1.01) and effective sample sizes. From the posterior distributions, we extracted fixed-effect coefficients for each OCM code and summarized them using posterior medians, 80% and 90% highest posterior density intervals (HPDIs), and the probability of direction (pd)—the proportion of the posterior distribution with the same sign as the median.

To visualize the results, for each outcome we selected OCM codes with 90% credible intervals excluding zero and plotted their posterior distributions (Figures 4 and 5). These exploratory analyses identify the cultural and behavioral domains most strongly associated with each punishment type, providing additional, ethnography-driven insight into the socio-behavioral contexts in which different forms of punishment are embedded.

Priors, computation, and reporting

Models were estimated with Hamiltonian Monte Carlo. We used weakly regularizing priors: $\mathcal{N}(0,1)$ for population intercepts and fixed-effect coefficients, and $Exponential(1)$ for random-effect standard deviations. Models were fit with four chains, sufficient iterations to achieve stable effective sample sizes (8,000 iterations), and target acceptance rates of

0.95–0.99 depending on the model. Convergence was assessed using standard diagnostics (trace plots, effective sample sizes, and RHAT near 1). We summarize posterior uncertainty using posterior medians and highest posterior density intervals (HPDIs). For preregistered hypothesis evaluation, we classified associations as supported when the 90% HPDI excluded zero in the predicted direction, moderately supported when the 80% HPDI excluded zero in the predicted direction, and not supported otherwise. We also report a posterior-direction metric (pd), defined as the proportion of posterior draws sharing the sign of the posterior median, as a continuous summary of directional evidence.

Data availability

Data are publicly available via the `ehrafpunishment`[CITE] R package, available on GitHub (https://github.com/zhgarfield/ehrafpunishmentdata) and archived on Zenodo [CITE]. NEED TO MAKE PUBLIC. Data to compute the TSD and RI factors are available in a fork of the `complex_coev_sccs` pacakge (https://github.com/zhgarfield/complex_coev_sccs).

Acknowledgements

ZHG, CM, and FDP acknowledge IAST funding from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program) and funding through an IAST Multidisciplinary Prize (award year: 2020). CM further acknowledges the support of the French Agence Nationale de la Recherche (ANR) under grant ANR-25-CE28-1200 (project ENFORCE). Alberto Michelletti contributed to the development and structure of this project.

References

1. Hill, K., Barton, M. & Hurtado, A. M. The emergence of human uniqueness: Characters underlying behavioral modernity. *Evolutionary Anthropology: Issues, News, and Reviews* **18**, 187–200 (2009).

2. Morin, E., Bird, D., Winterhalder, B. & Bliege Bird, R. Why Do Humans Hunt Cooperatively?: Ethnohistoric Data Reveal the Contexts, Advantages, and Evolutionary Importance of Communal Hunting. *Current Anthropology* **65**, 876–921 (2024).

3. Boyd, R. & Richerson, P. J. Large-scale cooperation in small-scale foraging societies. *Evolutionary Anthropology: Issues, News, and Reviews* **n/a**, (2022).

4. Ostrom, E. *Governing The Commons: The Evolution of Institutions for Collective Action*. (1990).

5. Glowacki, L., Wilson, M. L. & Wrangham, R. W. The evolutionary anthropology of war. *Journal of Economic Behavior & Organization* **178**, 963–982 (2020).

6. Wiessner, P. Collective Action for War and for Peace: A Case Study among the Enga of Papua New Guinea. *Current Anthropology* **60**, 224–244 (2019).

7. Zefferman, M. R. & Mathew, S. An evolutionary theory of large-scale human warfare: Group-structured cultural selection. *Evolutionary Anthropology: Issues, News, and Reviews* **24**, 50–61 (2015).

8. Balliet, D., Mulder, L. B. & Van Lange, P. A. M. Reward, punishment, and cooperation: A meta-analysis. *Psychological Bulletin* **137**, 594–615 (2011).

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

10. Jin, S., Spadaro, G. & Balliet, D. Institutions and cooperation: A meta-analysis of structural features in social dilemmas. *Journal of Personality and Social Psychology* **129**, 286–312 (2025).

11. Raihani, N. J. & Bshary, R. Punishment: One tool, many uses. *Evolutionary Human Sciences* **1**, 1–26 (2019).

12. Boehm, C. Purposive social selection and the evolution of human altruism. *Cross-Cultural Research* **42**, 319–354 (2008).

13. Fitouchi, L. & Singh, M. Punitive justice serves to restore reciprocal cooperation in three small-scale societies. *Evolution and Human Behavior* **44**, 502–514 (2023).

14. Henrich, J. Cooperation, Punishment, and the Evolution of Human Institutions. *Science* **312**, 60–61 (2006).

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

16. Baumard, N. Has punishment played a role in the evolution of cooperation? A critical review. *Mind and Society* 1–22 (2010) doi:10.1007/s11299-010-0079-9.

17. Balafoutas, L., Grechenig, K. & Nikiforakis, N. Third-party punishment and counter-punishment in one-shot interactions. *Economics Letters* **122**, 308–310 (2014).

18. Nikiforakis, N. & Engelmann, D. Altruistic punishment and the threat of feuds. *Journal of Economic Behavior & Organization* **78**, 319–332 (2011).

19. Herrmann, B., Thöni, C. & Gächter, S. Antisocial Punishment Across Societies. *Science* **319**, 1362–1367 (2008).

20. Garfield, Z. H., Buckner, W., Wrangham, R. & Glowacki, L. Norm violation and punishment across human societies. (2022) doi:10.17605/OSF.IO/4VSNW.

21. Marlowe, F. W. *et al.* More 'altruistic' punishment in larger societies. *Proceedings of the Royal Society B: Biological Sciences* **275**, 587–592 (2008).

22. Molho, C., De Petrillo, F., Garfield, Z. H. & Slewe, S. Cross-societal variation in norm enforcement systems. *Philosophical Transactions of the Royal Society B: Biological Sciences* **379**, 20230034 (2024).

23. Molho, C., Tybur, J. M., Van Lange, P. A. M. & Balliet, D. Direct and indirect punishment of norm violations in daily life. *Nature Communications* **11**, 3432 (2020).

24. Garfield, Z. H., Syme, K. L. & Hagen, E. H. Universal and variable leadership dimensions across human societies. *Evolution and Human Behavior* **41**, 397–414 (2020).

25. Fehr, E. & Fischbacher, U. Third-party punishment and social norms. *Evolution and Human Behavior* **25**, 63–87 (2004).

26. Henrich, J. *et al.* Costly punishment across human societies. *Science* **312**, 1767–70 (2006).

27. Marlowe, F. W. *et al.* The 'spiteful' origins of human cooperation. *Proceedings of the Royal Society B: Biological Sciences* **278**, 2159–2164 (2011).

28. Singh, M. & Garfield, Z. H. Evidence for third-party mediation but not punishment in Mentawai justice. *Nature Human Behaviour* **6**, 930–940 (2022).

29. Barclay, P. Reputational benefits for altruistic punishment. *Evolution and Human Behavior* **27**, 325–344 (2006).

30. Fehr, E. & Gächter, S. Cooperation and Punishment in Public Goods Experiments. *American Economic Review* **90**, 980–994 (2000).

31. Jordan, J. J., Hoffman, M., Bloom, P. & Rand, D. G. Third-party punishment as a costly signal of trustworthiness. *Nature* **530**, 473–476 (2016).

32. Earp, B. D., McLoughlin, K. L., Monrad, J. T., Clark, M. S. & Crockett, M. J. How social relationships shape moral wrongness judgments. *Nature Communications* **12**, 5776 (2021).

33. Bowles, S. & Gintis, H. The Origins of Human Cooperation. 18 (2002).

34. Gintis, H. Strong Reciprocity and Human Sociality. *Journal of Theoretical Biology* **206**, 169–179 (2000).

35. Hamilton, W. D. The genetical evolution of social behaviour. I. *Journal of theoretical biology* **7**, 1–16 (1964).

36. Trivers, R. L. The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology* **46**, 35–57 (1971).

37. Guala, F. Reciprocity: Weak or strong? What punishment experiments do (and do not) demonstrate. *Behavioral and brain sciences* **35**, 1–15 (2012).

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

39. McCullough, M. E., Kurzban, R. & Tabak, B. A. Cognitive systems for revenge and forgiveness. *Behavioral and Brain Sciences* **36**, 1–15 (2013).

40. Hutcherson, C. A. & Gross, J. J. The moral emotions: A social–functionalist account of anger, disgust, and contempt. *Journal of Personality and Social Psychology* **100**, 719–737 (2011).

41. Tripp, T. M. & Bies, R. J. "Righteous" anger and revenge in the workplace: The fantasies, the feuds, the forgiveness. in *International handbook of anger: Constituent and concomitant biological, psychological, and social processes* 413–431 (Springer Science + Business Media, 2010). doi:10.1007/978-0-387-89676-2_24.

42. Molho, C., Tybur, J. M., Güler, E., Balliet, D. & Hofmann, W. Disgust and Anger Relate to Different Aggressive Responses to Moral Violations. *Psychological Science* **28**, 609–619 (2017).

43. Seip, E. C., Van Dijk, W. W. & Rotteveel, M. Anger motivates costly punishment of unfair behavior. *Motivation and Emotion* **38**, 578–588 (2014).

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

45. Mathew, S. & Boyd, R. Punishment sustains large-scale cooperation in prestate warfare. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 11375–80 (2011).

46. Hames, R. Kin selection. in *The Handbook of Evolutionary Psychology* (ed. Buss, D.) vol. 2 505–523 (John Wiley & Sons, Inc., 2015).

47. Pedersen, E. J. *et al*. When and Why Do Third Parties Punish Outside of the Lab? A Cross-Cultural Recall Study. *Social Psychological and Personality Science* **11**, 846–853 (2020).

48. Lopez, L. D., Moorman, K., Schneider, S., Baker, M. N. & Holbrook, C. Morality is relative: Anger, disgust, and aggression as contingent responses to sibling versus acquaintance harm. *Emotion* **21**, 376–390 (2021).

49. McManus, R. M., Kleiman-Weiner, M. & Young, L. What We Owe to Family: The Impact of Special Obligations on Moral Judgment. *Psychological Science* **31**, 227–242 (2020).

50. Enke, B. Kinship, Cooperation, and the Evolution of Moral Systems. *The Quarterly Journal of Economics* **134**, 953–1019 (2019).

51. Henrich, J. *The Weirdest People in the World: How the West Became Psychologically Peculiar and Particularly Prosperous*. (Penguin UK, 2020).

52. Aktipis, A. *et al.* Understanding cooperation through fitness interdependence. *Nature Human Behaviour* **2**, 429–431 (2018).

53. Pleasant, A. & Barclay, P. Mutual cooperation gives you a stake in your partner's welfare, especially if they are irreplaceable. *Journal of Personality and Social Psychology* **129**, 1151–1162 (2025).

54. Krasnow, M. M., Delton, A. W., Cosmides, L. & Tooby, J. Looking Under the Hood of Third-Party Punishment Reveals Design for Personal Benefit. *Psychological Science* **27**, 405–418 (2016).

55. Henrich, J. *et al.* Markets, religion, community size, and the evolution of fairness and punishment. *Science* **327**, 1480–1484 (2010).

56. Schwartz, R. D. & Miller, J. C. Legal Evolution and Societal Complexity. *American Journal of Sociology* **70**, 159–169 (1964).

57. Turchin, P. *et al.* Quantitative historical analysis uncovers a single dimension of complexity that structures global variation in human social organization. *Proceedings of the National Academy of Sciences* **115**, E144–E151 (2018).

58. Fisher, R. J. Social Desirability Bias and the Validity of Indirect Questioning. *Journal of Consumer Research* **20**, 303–315 (1993).

59. Murphy, J. J., Allen, P. G., Stevens, T. H. & Weatherhead, D. A Meta-analysis of Hypothetical Bias in Stated Preference Valuation. *Environmental and Resource Economics* **30**, 313–325 (2005).

60. Wiessner, P. The role of third parties in norm enforcement in customary courts among the Enga of Papua. *Proceedings of the National Academy of Sciences* **117**, 32320–32328 (2020).

61. Ringen, E., Martin, J. S. & Jaeggi, A. Novel phylogenetic methods reveal that resource-use intensification drives the evolution of 'Complex' societies. (2021).

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

63. Naroll, R. The proposed HRAF probability sample. *Cross-Cultural Research* **2**, 70–80 (1967).

64. Subjects Covered (OCM) Human Relations Area Files.

65. Wiessner, P. Norm enforcement among the Ju/'hoansi Bushmen. *Human Nature* **16**, 115–145 (2005).

66. Blurton Jones, N. A selfish origin for human food sharing: Tolerated theft. *Ethology and Sociobiology* **5**, 1–3 (1984).

67. De Coccola, R., King, P. & Houston, J. *The incredible Eskimo: Life among the barren land Eskimo*. (Hancock House, 1986).

68. Oberg, K. Crime and punishment in Tlingit society. *American Anthropologist* **Vol. 36**, 145–155 (1934).

69. Turnbull, C. M. The Mbuti Pygmies: An ethnographic survey. *New York* (1965).

70. Gusinde, M. & Schütze, F. The Fireland Indians: Vol. 1. The Selk'nam, on the life and thought of a hunting people of the Great Island of Tierra del Fuego. in *Expeditions* 32, 1176, 51 plates (Verlag der Internationalen Zeitschrift, 1931).

71. Boyd, R., Gintis, H. & Bowles, S. Coordinated Punishment of Defectors Sustains Cooperation and Can Proliferate When Rare. *Science* **328**, 617–620 (2010).

72. Boehm, C. Costs and benefits in hunter-gatherer punishment. *The Behavioral and Brain Sciences* **35**, 19–20 (2012).

73. Emmons, G. Thornton. & De Laguna, F. The Tlingit Indians. in *Anthropological papers of the American Museum of Natural History* xl, 488 (University of Washington Press ; American Museum of Natural History, 1991).

74. Downes, R. M. The Tiv tribe. in (The Government Printer, 1933).

75. Price, S. *Co-wives and Calabashes*. (Univ of Michigan Press, 1993).

76. Ames, D. W. Belief in 'witches' among the rural Wolof of the Gambia. *Africa* **Vol. 24**, 263–273 (1959).

77. Peters, E. L., Goody, J. & Marx, E. The Bedouin of Cyrenaica: Studies in personal and corporate power. in *Cambridge studies in social and cultural anthropology* xii, 310 (Press Syndicate of the University of Cambridge, 1990).

78. Abu-Lughod, L. *Writing women's worlds: Bedouin stories*. (University of California Press, 1993).

79. Masters, W. M. *Rowanduz: A Kurdish administrative and mercantile center*. ([s.n.], 1953).

80. Leach, E. R. *Political systems of highland Burma: A study of Kachin social structure*. (Harvard University Press, 1954).

81. Hulbert, H. B. The passing of Korea. in (Doubleday, Page; Co., 1906).

82. DeGlopper, D. R. Religion and ritual in Lukang. (1974).

83. Donner, W. The first face. in *The five faces of Thailand : An economic geography, by Wolf Donner* 209–399, 913–922 (St. Martin's Press, 1978).

84. Rattray, R. S. (Robert. S. Ashanti law and constitution. in (Clarendon Press, 1929).

85. Pierson, D. & Lantz, H. R. Negroes in Brazil. in *Perspectives in sociology* lxxxiii, 420 (Southern Illinois University Press, 1967).

86. Garfield, Z. H., von Rueden, C. & Hagen, E. H. The evolutionary anthropology of political leadership. *The Leadership Quarterly* **30**, 59–80 (2019).

87. Boehm, C. *Hierarchy in the forest: The evolution of egalitarian behavior*. (Harvard University Press Cambridge, MA, 1999).

88. Singh, M., Wrangham, R. W. & Glowacki, L. Self-interest and the design of rules. *Human Nature* **28**, 457–480 (2017).

89. North, D. C. *Institutions, Institutional Change and Economic Performance*. (Cambridge University Press, 1990). doi:10.1017/CBO9780511808678.

90. Robinson, J. & Acemoglu, D. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. (Random House, 2012).

91. Fiske, A. P. The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological review* **99**, 689 (1992).

92. Rai, T. S. & Fiske, A. P. Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review* **118**, 57–75 (2011).

93. Watts, J. *et al.* Building Quantitative Cross-Cultural Databases From Ethnographic Records: Promise, Problems and Principles. *Cross-Cultural Research* **56**, 62–94 (2022).

94. Slingerland, E. *et al.* Coding culture: Challenges and recommendations for comparative cultural databases. *Evolutionary Human Sciences* **2**, e29 (2020).

95. Purzycki, B. G., Bendixen, T. & Lightner, A. D. Coding, causality, and statistical craft: The emergence and evolutionary drivers of moralistic supernatural punishment remain unresolved. *Religion, Brain & Behavior* **13**, 207–214 (2023).

96. Garfield, Z. H., Rueden, C. R. von & Hagen, E. H. The Multi-Capital Leadership Theory. *Human Nature* (2025) doi:10.1007/s12110-025-09503-y.

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

98. Gabry, J. & Češnovar, R. *Cmdstanr: R interface to CmdStan*. (2021).

99. Murdock, G. P. & White, D. R. Standard cross-cultural sample. *Ethnology* **8**, 329–369 (1969).

100. Hammarström, H., Forkel, R., Haspelmath, M. & Bank, S. Glottolog 5.1. *Leipzig: Max Planck Institute for Evolutionary Anthropology* **10**, (2024).

101. Piironen, J. & Vehtari, A. Sparsity information and regularization in the horseshoe and other shrinkage priors. *Electronic Journal of Statistics* **11**, 5018–5051 (2017).

102. Carvalho, C. M., Polson, N. G. & Scott, J. G. Handling sparsity via the horseshoe. in *Proceedings of the twelfth international conference on artificial intelligence and statistics* (eds. Dyk, D. van & Welling, M.) vol. 5 73–80 (PMLR, 2009).