The Cognitive and Cultural Foundations of Moral Behavior[☆]

Benjamin Grant Purzycki[a,*], Anne C. Pisor[a], Coren Apicella[b], Quentin Atkinson[c,d], Emma Cohen[e,f], Joseph Henrich[g], Richard McElreath[a], Rita A. McNamara[h], Ara Norenzayan[i], Aiyana K. Willard[e], Dimitris Xygalatas[j]

[a]*Department of Human Behavior, Ecology, and Culture, Max Planck Institute for Evolutionary Anthropology, DE*
[b]*Department of Psychology, University of Pennsylvania, USA*
[c]*Department of Psychology, University of Auckland, NZ*
[d]*Max Planck Institute for the Science of Human History, DE*
[e]*Institute of Cognitive and Evolutionary Anthropology, University of Oxford, UK*
[f]*Wadham College, University of Oxford, UK*
[g]*Department of Human Evolutionary Biology, Harvard University, USA*
[h]*School of Psychology, Victoria University of Wellington, NZ*
[i]*Department of Psychology, University of British Columbia, CA*
[j]*Department of Anthropology, University of Connecticut, USA*

Abstract

Does moral culture contribute to the evolution of cooperation? Here, we examine individuals' and communities' models of what it means to be good and bad and how they correspond to corollary behavior across a variety of socioecological contexts. Our sample includes over 600 people from eight different field sites that include foragers, horticulturalists, herders, and the fully market-reliant. We first examine the universals and particulars of explicit moral models. We then use these moral models to assess their role in the outcome of an economic experiment designed to detect systematic, dishonest rule-breaking favoritism. We show that individuals are slightly more inclined to play by the rules when their moral models include the task-relevant virtues of "honesty" and "dishonesty." We also find that religious beliefs are better predictors of honest play than these virtues. The predictive power of these values' and beliefs' local prevalence, however, remains inconclusive. In summary, we find that religious beliefs and moral models may help promote honest behavior that may widen the breadth of human cooperation.

Keywords: morality, cross-cultural ethnography, cognitive anthropology, evolution of cooperation

1. Introduction

Many theories hold that socially learned moral norms are the lynchpin for the remarkable breadth of cooperation that humans uniquely exhibit (Bowles and Gintis, 2003; Boyd, 2017; Boyd and Richerson, 2009; Chudek and Henrich, 2011; Richerson et al., 2016). However, there are a few critical outstanding issues that make this view difficult to endorse with a confidence borne out by direct empirical evidence. First, it is not immediately obvious that individuals' and groups' moral

[☆]Original Article

[*]Corresponding author
Email address: benjamin_purzycki@eva.mpg.de (Benjamin Grant Purzycki)

prescriptions actually influence the behavior of those who espouse them (e.g., Graham et al., 2015; Haidt, 2001; Perry, 2017; Smith et al., 2013). When moral prescriptions and behavior are consistent with each other, moral prescriptions might simply be rationalizations of behavior rather than causes (e.g., Baumard, 2016; Haidt, 2001). Second, despite the fact that so many emphasize (or minimize) the importance of culture for human cooperation, few actually measure its effects directly and model it as a distributed, superordinate property of social life (see Smaldino, 2014). Most empirical studies consider culture indirectly by either: a) having participants in economic experiments make an allocation with money and then asking what the appropriate decision was (e.g., Gurven et al., 2008; Ensminger and Henrich, 2014; Henrich and Henrich, 2014), b) framing experimental introductions in locally salient ways (e.g., Brodbeck et al., 2013; Cohn et al., 2014; Cronk, 2007; Lesorogol, 2007; Gerkey, 2013; Stagnaro et al., 2017), or c) conducting studies across multiple groups, and concluding that cross-cultural variation in behavior reflects underlying variation in culture (e.g., Apicella et al., 2012; Henrich, 2000; Henrich et al., 2004; Roth et al., 1991). Third, many cross-cultural studies emphasizing the evolved psychology underlying morality rely heavily on theoretically-motivated scale designs (e.g., Curry et al., n.d.; Graham et al., 2011) that a) use items lacking in local relevance, b) are impractical for innumerate and/or nonliterate populations, c) presuppose that samples have the lexical equivalent of "moral," and d) do not link this data to quantitative behavior.

Here, we seek to overcome these limitations by measuring moral culture from a variety of societies and examine whether or not moral values and their distributions actually have an impact on the kind of broader cooperation typified by humans. We first briefly spell out our assumptions and introduce contemporary evolutionary perspectives on moral systems, followed by a more detailed assessment of the aforementioned limitations. We then introduce our two studies. The first consists of an analysis of systematically collected ethnographic data regarding what it means to be "good" and "bad" across eight different field sites. In doing so, we examine cross-cultural moral universals and local particulars. The second study uses this data to examine its contribution to corresponding behavior in an experimental game designed to distinguish dishonest favoritism from impartial, rule-following fairness. We conclude with a discussion of our studies' limitations and comment on avenues for further inquiry.

2. Background

2.1. Defining moral systems

We refer to "moral models" here as the content and structure of individuals' explicit representations of moral norms. If we adopt the view that "culture" is shared, socially transmitted information (cf. Boyd and Richerson, 1988; D'Andrade, 1981; Sperber, 1996), then moral culture is the shared, socially transmitted units that comprise individual moral models. Defined in this fashion, local prevalence of particular units of socially transmitted information indicates how "cultural" or "normative" those units are. In this view, then, directly assessing whether or not culture influences individual behavior requires: 1) detailing individuals' models, 2) assessing how widespread the content of those models is in individuals' social groups, 3) examining the relationship between a behavioral trait and an individuals' models, and 4) examining the relationship between the trait and how prevalent specific informational units are in one's group. The first two requirements are descriptive, ethnographic accounts of moral culture. The latter most two allow us to disambiguate the relative impacts of individual and cultural models of morality on behavior. If moral culture predicts moral behavior, then the prevalence of moral models' constituent units in a group should covary with the target behavior.

We use "moral systems" here to refer to moral models, their psychological underpinnings, behavioral expressions, cultural prevalence, and the causal links between them (cf. Alexander, 1987; Haidt and Kesebir, 2010; Kiper and Sosis, 2014). Classical philosophical and contemporary social psychological views of moral systems emphasize universality and/or the view that morality is associated with abstract notions like "justice" and "rights" (Caton, 1963; Kant, 1997 [1785]; Turiel, 1983, 2006). In contrast, many evolutionary views boil down moral systems to the regulation of cooperative and/or mutualistic endeavors that generate individual- and/or group-level benefits (Alexander, 1987; Baumard et al., 2013; Barrett et al., 2016; Cosmides and Tooby, 2005; Cronk, 1994; Curry, 2016; Darwin, 1871; Greene, 2013; Fehr et al., 2002; Haidt and Kesebir, 2010; Machery and Mallon, 2010; Mizzoni, 2009; Sripada and Stich, 2006; Trivers, 1971; Tomasello and Vaish, 2013). However, there is considerable variation in moral systems, variation that many suggest are inconsequential or run counter to such generalist theories (Baumard, 2016; Boehm, 1980; Buchtel et al., 2015; Fessler et al., 2015; Schwartz, 2007; Shweder et al., 1997; Smith et al., 2007). As we detail below, piecing together the constituent parts of moral systems in a cross-cultural empirical project remains a major challenge in the evolutionary literature.

2.2. Measuring components of moral systems

2.2.1. Evolutionary psychology of morality

Contemporary evolutionary psychological research focused on mapping the conceptual space of morality typically relies on scale items (Curry et al., n.d.; Graham et al., 2011) with prefabricated materials that are verified externally (i.e., using other scales). For example, seeking to better operationalize the moral domain with attention to cross-cultural validity, the popular "Moral Foundations" literature breaks down the evolutionary and cognitive "foundations" of morality into a few core dimensions. While the rubric itself has evolved (Graham et al., 2013), the most recent iteration includes: (1) harm/care; (2) fairness/reciprocity; (3) ingroup/loyalty; (4) authority/respect, and (5) purity/sanctity as foundational to moral reasoning. The more recent "Morality-as-Cooperation" literature (Curry, 2016; Curry et al., n.d.) measures seven types of cooperation treated as the foundations for moral behavior: (1) family values; (2) group loyalty; (3) reciprocity; (4) dominance; (5) deference; (6) fairness; and (7) rights to property.

These rubrics were not designed to assess the relationship between moral culture and behavior. Rather, they seek to identify variation in moral reasoning as indicated by variation in how survey items load onto principal components and how mean values of scales vary across different groups. There are practical and methodological reasons to be reluctant to employ scale-based surveys in populations where they were not designed. First, many traditions lack the lexical equivalent of "morality." Second, some samples struggle with scale-based survey instruments. While convenient for researchers, in practice, scale items can be quite taxing and unintuitive for non-literate and/or innumerate participants (e.g., Gurven et al., 2013). Third, such instruments are often limited in local relevance. For example, the "Moral Foundations Questionnaire" (Graham et al., 2011) includes questions about whether or not "being good at math," having "love for one's country," being "denied rights," and "God's approval" are "relevant to [participants'] moral thinking" or to their sense of right and wrong. Such items and the notion of "moral relevance" are simply unintelligible in many contexts. Ideally, scale design in cross-cultural research begins with preliminary ethnographic inquiry to ensure that scale items are actually measuring target constructs (Bernard, 2011; Handwerker, 2001). Indeed, Smith et al. (2007) found that other theory-driven classification schemes inadequately captured the variation in folk-models of what it means to be "good" in seven

different communities. Boehm (1980) imported a morality metric to Montenegro, but due to participants' initial off-target responses to the metric, he had to assess features of local moral behavior with open-ended questions.

2.2.2. Cultural evolutionary ecology of moral behavior

Those who emphasize culture's effects on cooperative behavior typically employ economic experimental games as an index of cooperation, but do not directly measure or model "culture." Some appeal to the importance of cultural institutions (i.e., shared pools of norms that constrain human interactions in specific, socially demarcated contexts; see D'Andrade 2006; North 1991; Searle 1995) by manipulating the cultural relevance of experiments' instructions in the form of framing effects (Brodbeck et al., 2013; Cohn et al., 2014; Cronk, 2007; Lesorogol, 2007; Gerkey, 2013). Others conduct experiments and infer that culture contributes to the evolution of cooperation by virtue of statistical divergences between groups in experimental game outcomes (Apicella et al., 2012; Ensminger and Henrich, 2014; Henrich, 2000; Henrich et al., 2004; Roth et al., 1991). A burgeoning literature that actively measures variation in cultural information focuses on religious beliefs (McNamara et al., 2016; Johnson, 2016; Purzycki et al., 2016b). This literature typically uses individuals' beliefs in punitive and knowledgeable deities to predict cooperative outcomes. However, the literature ignores the within-group distribution of religious beliefs–that is, groups' religious culture–as a factor in individual behavior.

To the best of our knowledge, only a solitary study claims to assess the degree to which specifically *moral* culture has an effect in cooperation (Gurven, Zanolini, and Schniter, 2008; cf. Ensminger and Henrich, 2014; Henrich and Henrich, 2014 for more case studies from the same cross-cultural project). The study, which the authors characterize as "the first of its kind to show that local culture matters in explaining variation in pro-social behavior" (pg. 589), employed a variation of the Dictator Game in which participants were asked to identify "the morally correct offer to give in this...game" (3). While there were a few exceptions in their sample of nine Tsimane' (Bolivia) villages, in general there was a positive correlation between what people thought they should do in the game and how they actually allocated money. Note that the Tsimane' lack a word for "moral"; the researchers instead used *ruijsis* which "expresses the concept of appropriate behavior or action" (pg. 592). The authors "argue that local differences [in cooperation] are not necessarily due to strong norms *per se* that vary among villages, but due to local (unmeasured) effects that push and pull villages towards more or less pro-social sentiment" (pg. 589). Note, however, that this study elicited participants' views of the "appropriate" behavior for the experiment, not more general individual-level moral models or their distributions from which participants ostensibly drew.

In summary, while many argue about culture's role in the expansion of cooperation, and many examine a variety of important factors' contributions to this process, the cooperation literature does not directly probe the contribution of moral culture itself. Likewise, studies that have measured variation in moral culture do not link them to corresponding behavior. Here, we attempt to assess the relationship between individual moral models and cooperation directly, with an eye to the aspects of local culture—moral models' prevalence—that may serve as inputs to individual behavior. First, we assess the degree to which there are universals and particulars of moral culture by examining what people consider "good" and "bad" (Study 1). We then examine the respective roles of individual moral models and culture in the allocation of money using an experimental economic game designed to measure honest, impartial rule-following behavior toward anonymous others (Study 2).

3. Participants

We collected data in eight different field sites (see table 1). These samples included: (1) the Hadza of Tanzania; (2 and 3) inland and coastal villagers from Tanna, Vanuatu; (4) residents of Marajó island in Brazil; (5) Fijians from Yasawa island; (6) Indo-Fijians from Lovu; (7) Tyvan residents in Kyzyl, Tyva Republic; and (8) Indo-Mauritian residents of Porte aux Piment. Our sample is notably diverse; modes of subsistence range from the foraging Hadza and horticultural inland Tannese to the fully market-integrated economies of Kyzyl and Porte aux Piment.

Site/Sample	Researcher	Sampling Method	World Religion	Economy
Coastal Tanna	Atkinson	Cluster sampling (census)	Christianity	Horticulture/Market
Hadza	Apicella	Entire camps	None	Foraging
Inland Tanna	Atkinson	Entire community	None	Horticulture
Lovu, Fiji	Willard	Door-to-door	Hinduism	Market
Mauritians	Xygalatas	Random (street) sampling	Hinduism	Farming/Market
Marajó, Brazil	Cohen	Random sampling (census)	Christianity	Market
Tyva Republic	Purzycki	Random and chain sampling (street)	Buddhism	Herding/Market
Yasawa, Fiji	McNamara	Door-to-door	Christianity	Horticulture/Market

Table 1: **Descriptive features of each field site.** See Purzycki et al. (2016a) and supplements for further details.

This sample exhibits some of the considerable cross-cultural diversity known to cultural anthropology; our participants range from the fully market-integrated (e.g., Mauritians and Marajó Brazilians) to subsistence foragers and horticulturalists (e.g., the Hadza and Inland Tanna, respectively). Our sample thus includes people from traditional, small-scale communities, whose means of living are subsistence-based with daily interactions consisting primarily with local familiars as well as urban samples where individuals regularly interact with anonymous others. Moreover, our samples are uniquely poised for consideration of our research questions and methods. No such work examining explicit moral models has been conducted among traditional, small-scale societies. Our methods (see below) are particularly useful for eliciting rich and comparable ethnographic data in innumerate and nonliterate samples often ignored or left out in studies relying on prefabricated scales and narrow samples.

4. Study 1: Moral Culture

While not every group has the lexical equivalent of "moral," some posit that the distinction between "good" and "bad" is a human universal (Brown, 1991; Wierzbicka, 1994). We can assess whether or not the content of these conceptual domains approximate morality simply by asking people what it *means* to be good and bad. To reliably capture moral models and culture, we assess freely-elicited data of what it means to be "good" and "bad" (see Buchtel et al., 2015; Purzycki, 2011, 2016; Smith et al., 2007, for precedent applications). This method allows individuals to answer on their own terms and avoids the aforementioned pitfalls associated with a lack of cultural relevance, the question of what measurement instruments actually measure, or the elicitation of rationalizations of behavior.

4.1. Methods

All materials were translated in local languages and back-translated into English for corroboration and subsequent edits. To obtain reliable, naturalistic, and culturally relevant data about morality, we asked participants to:

- *Please list up to 5 behaviors that make someone a good/virtuous/moral person.*
- *Please list up to 5 behaviors that make a bad/immoral person.*

All free-list data were translated into English and subsequently submitted to Purzycki for compiling and coding. All original open-ended responses in English and the subsequently coded data are publicly available here `https://github.com/bgpurzycki/Moral-Models-Moral-Behavior` for reassessment, further recoding, and analysis. See supplements for further materials, methodological notes, and per-site English translations of instructions of the free-list tasks.

4.1.1. Analysis

We analyzed the free-list data using the AnthroTools package (Jamieson-Lane and Purzycki, 2016; Purzycki and Jamieson-Lane, 2016) for R (R Core Team, 2016). This package calculates the cognitive salience of individual free-list items and tabulates their mean salience score (Smith's S). These scores can be calculated at the sample- and sub-sample (i.e., field site) levels. Individual item salience (i) is calculated with equation 1:

$$i = \frac{n + 1 - k}{n} \tag{1}$$

where n is the total number of items an individual listed, and k is the order in which an item was listed. Smith's S (equation 2) is a sample's mean value of item type (Smith, 1993; Smith et al., 1995; Smith and Borgatti, 1997):

$$\mathrm{S} = \frac{\sum i_T}{N} \tag{2}$$

Here, we denote item type with i_T and N denotes the total sample or sub-sample size (i.e., the denominator is not limited only to those participants who listed a given item). Smith's S will therefore increase as a function of ubiquity and earlier placement in lists. In order to minimize inflated Smith's S values due to repeated items within lists, we used AnthroTools' "MAX" function which includes only the earliest-listed repeated item in its calculations.

It is important to note that our reported salience indices may reflect underestimations for three reasons. First, in terms of individual item salience, participants were encouraged to list up to only 5 items per sub-domain (i.e., "good" and "bad"). If these items would have been the earliest-listed in a task without such a constraint, all data considered here would have had much higher salience scores. Secondly, we retain items listed by only one individual in our analysis. Dropping such idiosyncratic items would decrease the denominator in equation 1 and therefore increase the salience values. Thirdly, for the sake of completeness, we include those individuals who simply answered "I don't know" (1 participant for the "good" list and 3 for the "bad"). Though negligible, these would inevitably have an item salience of 1 and contribute to a larger denominator in equation 2.

4.2. Results

4.2.1. Moral Universals

What constitutes a "moral" or "good" person? Participants ($N = 643$) listed a total of 2,478 items ($M_{Listed} = 4.27$, $SD = 1.07$) in this sub-domain. Table 2 and Figure 1 show the salience of individually-listed items where $\mathrm{S} \geq 0.10$. We use this cut-off to minimize table lengths. Many participants listed various items that simply re-expressed the question (e.g., good people have "good hearts" or exhibit "good behavior"). For the purposes of analysis, these items were given the same

code. After this, the most salient item for participants was "generosity" or "sharing," followed by "helpfulness" and "honesty."

For the sake of reference and discussion, we include a *post hoc* coding of item types by their corresponding categories in the Moral Foundations (Graham et al., 2011, 2013) and Morality-as-Cooperation (Curry, 2016; Curry et al., n.d.) literatures. If we take these most salient items and apply their equivalent label in the Foundations and Cooperation typologies, it is clear that items in the "fairness/reciprocity" domain are the most salient. This suggests a greater cultural stability for this "foundational" category. While "honesty" appears in the Foundations typology and ranks among the items with the highest salience in the present free-lists, there is not a broad consensus about honesty in the literature; others simply include honesty as another moral subdomain (Ashton et al., 2014; Blasi, 1980; Hofmann et al., 2014) of the greater repertoire of moral foundations and the Morality-as-Cooperation literature bypasses it. Regardless, these results are consistent with what we would expect given the evolutionary literature's view of morality as a system regulating social exchange and cooperation (Alexander, 1987; Curry et al., n.d.): fairness and reciprocity loom large in mental models of what it means to be good. "Loving" and being kind are also included. While the Foundations literature considers this as part of the Care/Harm foundation, it is not immediately clear how these fall within the scope of the Cooperation typology, a limitation acknowledged in Curry et al. (n.d.).

Item	Foundation	Cooperation Type	Salience (M)	Smith's S	n
Good*	—	—	0.73	0.22	195
Generous/Shares	Fairness/Reciprocity	Reciprocity	0.64	0.18	178
Helpful	Fairness/Reciprocity	Reciprocity	0.59	0.17	183
Honest	Honesty/Deception	?	0.69	0.14	129
Respectful	Authority/Respect	Group Loyalty	0.56	0.13	144
Loving	Care/Harm	?	0.61	0.11	113
Kind	Care/Harm	?	0.73	0.10	88

Table 2: **Sample salience scores $\geq$ 0.10 for what makes a good person (N = 643).** Salience (M) is the average individual item salience among individuals who listed a given item, Smith's S is the individual item salience of the sample, and n is the number of participants listing the item. Foundation column indicates corresponding type in the "Moral Foundations" literature, while Cooperation Type column denotes corresponding type in the "Morality-as-Cooperation" literature. Question marks indicate possible interpretations or lack of obvious correspondence. *Denotes clustered items such as good conscience, good behavior, good nature, good heart.

What makes an "immoral" or "bad" person? In this sub-domain, participants ($N = 650$) listed 2,728 items ($M = 4.20$, $SD = 1.14$). Table 3 details the top five items where S $\geq$ 0.10. Figure 1 illustrates the content of these models by salience. Much like the "good" data, many participants reiterated the question in their responses (e.g., bad people exhibit "bad behavior"). Also consistent with the evolutionary literature, "theft" is the most salient item listed across cultural groups (the antithesis of generosity), followed by "deceit" and "violence." Notably—and not often considered in much of the evolutionary literature (cf. Kurzban et al., 2010; Rozin, 1999)—the use and abuse of drugs, alcohol, and other substances are among the chief items listed in what makes a "bad" person. It may be that people really do see the consumption of drugs and alcohol as bad in and of itself, but they also may view intoxicants as the source of bad behavior and therefore indicative of immoral conduct. Here too, the Foundations literature has a little more coverage than the Morality-as-Cooperation literature.

Item	Foundation	Cooperation Type	Salience (M)	Smith's S	n
Theft	Fairness/Reciprocity	Property Rights	0.71	0.26	235
Dishonest	Honesty/Deception	?	0.66	0.16	156
Violent	Care/Harm	?	0.65	0.15	153
Drugs/Alcohol	Purity/Sanctity (?)	?	0.70	0.12	108
Bad*	—	—	0.71	0.11	101

Table 3: **Sample salience scores $\geq$ 0.10 for what makes a bad person (N = 650).** Salience (M) is the average individual item salience among individuals who listed a given item, Smith's S is the individual item salience of the sample, and n is the number of participants listing the item. Foundation column indicates corresponding type in the "Moral Foundations" literature, while Cooperation Type column denotes corresponding type in the "Morality-as-Cooperation" literature. Question marks indicate possible interpretations or lack of obvious correspondence. *Denotes clustered items such as bad conscience, bad behavior, bad nature, bad heart.

Cross-culturally, the most salient components of individuals' mental models of morality revolve around the provisioning of material resources in the form of generosity, helpfulness, and theft. We might interpret honesty and dishonesty as facets of material goals as well insofar as it is virtuous to be honest about how much others stand to gain or lose in interactions (per Ashton et al., 2014; Fischbacher and Föllmi-Heusi, 2013, see below). However, as is made clearer in our examination of site-specific moral models, there is also some variation that is not immediately related to cooperation.

4.2.2. Moral Particulars

While we presently have neither the room to contextualize each site's specific results nor the means to make confident inferences about why cross-cultural variation exists (i.e., we only have 8 groups and cross-sectional data), the analysis of moral particulars does indicate that patterns in the global sample are not merely artifacts of a few groups driving the results. For the sake of concision, we present the most salient item listed by site. Table 4 details the items with the highest salience scores by site for the "good" free-list data. Table 5 details the items with the highest salience scores by site for the "bad" list. More exhaustive tables are available in the supplementary materials (Tables S1 and S2).

Site	Item	Foundation	Cooperation Type	Salience (M)	Smith's S	n
Coastal Tanna	Generous/Shares	Fairness/Reciprocity	Reciprocity	0.74	0.40	44
Hadza	Loving*	Care/Harm	?	0.62	0.40	69
Inland Tanna	Hospitable	Care/Harm	Group Loyalty (?)	0.74	0.45	74
Lovu, Fiji	Honest	Honesty/Deception	?	0.73	0.38	79
Mauritius	Speaks well	Authority/Respect (?)	Deference (?)	0.84	0.33	82
Marajó, Brazil	Helpful	Fairness/Reciprocity	Reciprocity	0.70	0.18	77
Tyva Republic	Honest*	Honesty/Deception	Reciprocity (?)	0.74	0.28	115
Yasawa, Fiji	Goes to church	Purity/Sanctity (?)	Group Loyalty (?)	0.68	0.39	103

Table 4: **Per-site items with highest salience scores for what makes a good person.** Salience (M) is the average individual item salience within individuals who listed a given item and Smith's S is the individual item salience of the sub-sample. Sample size (n) is site-specific sample size for sub-domain. Foundation column indicates corresponding type in the "Moral Foundations" literature, while Cooperation Type column denotes corresponding type in the "Morality-as-Cooperation" literature. Question marks indicate possible interpretations or lack of obvious correspondence. *Second highest salience within sample after the lumped "good" category (see above).

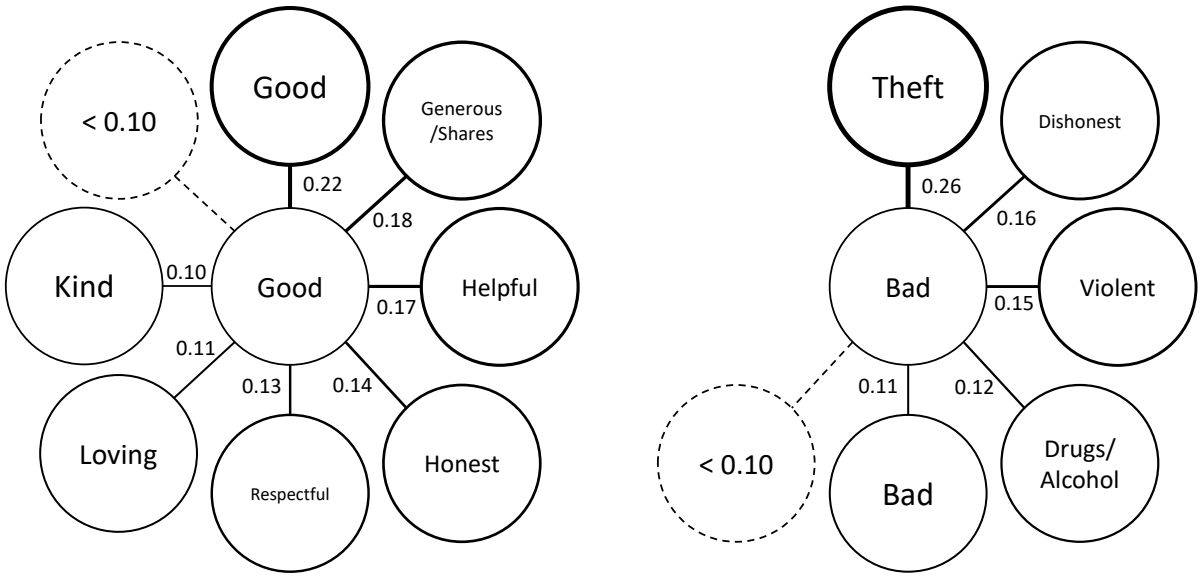

Figure 1: **Universal moral models**. Center circle represents the domain. The most salient items in these domains are on top with item connection weights (Smith's S values) descending clockwise.

Across the most salient items of what constitutes good people, we find some curious idiosyncracies. For instance, Indo-Mauritians listed "speaks well" (i.e., polite and considerate as opposed to rude and obscene) earlier and more often in lists than did other groups. Notably, Indo-Fijians (Lovu, Fiji) likewise ranked "speaks well" highly (S = 0.10) while all other sites ranked it lower when listed at all (S values were ≤ 0.04). This suggests a distinct cultural lineage; both are Indian diaspora populations that may have common-source value systems. Going to church was the most salient item for Fijians from Yasawa. This particular value does not obviously fit in the Moral Foundations or Morality-as-Cooperation rubrics and was not ranked highly at any other site. We nevertheless suggest "foundations" or "cooperation type" into which these might be classified in Tables 4 and 5. Likewise, in Marajó, "ignorance/arrogance" has the highest salience, although the small Smith's S suggests minimal consensus at this site. This term *ignorante* is locally nuanced, and refers to one who ignores others' opinions and holds their own as superior. This is also not obviously a component of the Moral Foundations rubric. Note that for half of the sites, "theft" had the highest salience for what constitutes a bad person; for the other half, save Marajó, group-level salience for theft was >0.10: Hadza (S = 0.31); Mauritius (S = 0.15); Tyva Republic (S = 0.13). Again, we refer readers to the more thorough tables in the supplementary materials.

4.2.3. Discussion

Cross-culturally, the most salient indicators of good people are generosity and sharing, helpfulness, and honesty. The most salient indicators of bad people are theft, dishonesty, and violence. Theft is notable insofar as it might be construed as the antithesis of cooperation; Baumard (2016) characterizes it as "violating the logic of balanced interests" (126). These items largely correspond to rubrics offered by the Foundations and Cooperation literatures, but a closer look at site-specific moral models appear to complicate those rubrics insofar as articulateness, religious piety, drug and alcohol use and abuse, and ignorance and arrogance are not consistently considered. Aside from il-

Site	Item	Foundation	Cooperation Type	Salience (M)	Smith's S	n
Coastal Tanna	Theft	Fairness/Reciprocity	Property Rights	0.76	0.46	44
Hadza	Murder	Care/Harm	?	0.52	0.34	69
Inland Tanna	Theft	Fairness/Reciprocity	Property Rights	0.76	0.37	78
Lovu, Fiji	Theft	Fairness/Reciprocity	Property Rights	0.82	0.36	79
Mauritius	Violence	Care/Harm	?	0.65	0.30	84
Marajó	Ignorance/Arrogance	Authority/Respect (?)	Deference (?)	0.69	0.14	76
Tyva Republic	Dishonest	Honesty/Deception	?	0.70	0.35	117
Yasawa, Fiji	Theft	Fairness/Reciprocity	Property Rights	0.80	0.33	103

Table 5: **Per-site items with highest salience scores for what makes a bad person.** Salience (M) is the average individual item salience within individuals who listed a given item and Smith's S is the individual item salience of the sub-sample. Sample size (n) is site-specific sample size for sub-domain. Foundation column indicates corresponding type in the "Moral Foundations" literature, while Cooperation Type column denotes corresponding type in the "Morality-as-Cooperation" literature. Question marks indicate possible interpretations or lack of obvious correspondence.

lustrating their limited breadth and narrow focus, our data pose little problem for the approaches; we are explicitly measuring representational models of morality rather than moral reasoning or judgment of a particular behavior. That said, it does suggest that the methods used in this literature have not captured some potentially informative variation found in cultural representations of what it means to be (im)moral.

While the above results map cross-cultural models of morality for eight populations, one remaining question is whether or not the content and ubiquity of moral models affect behavior. To assess this, we examined the effects of the presence and cognitive salience of honesty or dishonesty on a behavioral economic experiment that measured systematic and partial allocations. We predicted that in a game that measures dishonest favoritism indicative of systematic rule-breaking, the cognitive salience of "honesty" and "dishonesty" should predict fairer play. In other words, by measuring whether an individual's moral model includes honesty and dishonesty, we can assess his or her resistance to the opportunity to cheat. To assess the impact of moral culture, we also modeled group-level cultural prevalence of these components.

5. Study 2: Moral Behavior

5.1. Methods

5.1.1. Economic experiment

To measure honest behavior, we had participants play an economic game designed to detect dishonest favoritism (Cohn et al., 2014; Fischbacher and Föllmi-Heusi, 2013; Hruschka et al., 2014; Jiang, 2013; McNamara et al., 2016; Purzycki et al., 2016b). In this experiment, participants have a stack of 30 coins, a fair 6-sided, 2-colored die, and two cups designated for a specific individual. They think of which cup they want to put a coin into and then they roll the die. If one predesignated color appears, they are supposed to put the coin into the cup they thought of. If the other color appears, they are supposed to put the coin into the cup opposite the one they thought of. They repeat this until all 30 coins are in cups. They make these decisions alone without any outside observers. Regardless of individual decisions, coins should be randomly allocated to either cup if participants follow instructions. However, since participants play alone, they can allocate more coins to the cup of their preference.

In our study, participants played two counterbalanced games, each with two cup dyads. The "Local Community Game" included one cup reserved for an anonymous co-ethnic, co-religionist in the participant's local community and one was for an anonymous, geographically distant co-religionist who was also a co-ethnic by default. In the "Self Game," one cup was reserved for the player and the other cup for another anonymous co-ethnic, co-religious individual in the same specified geographically distant region. Players got to keep the money that went into their own cups and we distributed all allocations to randomly-selected individuals designated by the other cups. Here, the geographically distant players function as an index of the kind of broader cooperation that is unique to humans; as participants are not likely to ever interact with these distant players, playing according to the rules indicates an unwillingness to favor themselves or their community.

Show-up fees for participation were ∼25% of the average daily wage in our field sites. We set aggregate stakes at roughly a single day's wage (x) where individual die rolls were worth the closest coin in value to: x/number of games played/30 coins. Coins were real currency in each site except for the Hadza who played with tokens each worth 8 oz. (∼226.80g) maize. All participants were tested for game comprehension and knew that all coins would be distributed to those designated by the cups, including themselves. Only participants who passed the comprehension questions are included in this data set.

After experiments, participants answered a host of interview questions, including the aforementioned free-list tasks. Note that in experiments were predominantly those who completed free-list tasks, but there a few who did not complete free-lists or cases where individuals who participated in free-list tasks did not participate in experiments. They were also asked what they thought the experiment was about; their open-ended responses were coded for whether they mentioned cheating, fairness, or honesty. Participation took a total of ∼90 minutes, with the free-list task typically taking place ∼15 minutes after the game. Briefly, there are at least five reasons why this ordering had no effect on free-list outcomes: 1) free-list tasks were after demographic surveys, 2) the game check question is not correlated with listing (dis)honesty, 3) some sites simply did not list (dis)honesty frequently, 4) previous research (Smith et al., 2007) not using experiments in this fashion shows that listing "honest" is quite prevalent cross-culturally, and 5) comparing the data from one site with and without the experiment shows no indication of games having an effect (see Supplementary section 2.4 for more details).

All methods, materials, and data are available online at `https://github.com/bgpurzycki/Moral-Models-Moral-Behavior`.

5.1.2. Does moral culture matter?

If moral models contribute to the expansion of cooperation, listing (dis)honesty should predict playing by the rules. Participants who mention (dis)honesty will be more likely to allocate coins to the cup benefiting the recipient more socially distant from themselves—that is, someone non-local (as opposed to someone local), or someone other than the participant him- or herself. If the cognitive accessibility of task-relevant components of moral models is also important to the expansion of cooperation, we should see that salience of listing (dis)honesty increasing the chances of allocating coins as well. If moral culture contributes to the expansion of cooperation, then within-sample ubiquity of (dis)honesty in moral models should also have an effect on individual behavior. In other words, while an individual's moral model may induce fairer behavior, living in a context where more people share similar moral values–along with the expected repercussions of violating those values–should also contribute to the likelihood of playing honestly. Conversely, if

moral culture evolves in response to local problems, it may actually be associated with *more* self-interested behavior. Previous work we build upon shows that individual-level beliefs about morally concerned gods' punishment and knowledge breadth (i.e., omniscience) predicted allocations to the distant play (Purzycki et al., 2016b). By the same logic, then, the more one's community claims that morally concerned deities know and punish people, the more likely individuals should behave fairly (or not). In sum, in addition to individuals' moral models and religious beliefs, within-sample ubiquity of: (a) (dis)honesty in moral models, as well as (b) beliefs about moralistic gods' omniscience and (c) punishment, should predict fair play above and beyond the content of an.

To test whether moral models affect game play, we coded whether or not participants listed honesty or dishonesty in their free-lists and used the summation of these two indices as a predictor with possible values of 0, 1, and 2. As such, participants who did not answer free-list tasks or answered with "I don't know" are not considered in this analysis. Note that the odds of mentioning both honesty and dishonesty are related; a logistic regression shows that the odds of listing dishonesty increase by 7.10 [95% CI = 4.48, 11.25] when participants list honesty.

5.1.3. Participants

Table 6 reports the summary statistics for experimental participants who listed "honesty" in the "good" list ($n = 104$, 51 women, mean age = 37.26) or "dishonesty" in the bad list ($n = 130$, 49 women, mean age = 39.32). Note the considerable variability in the number of participants who listed honesty or dishonesty across these eight populations. Ten individuals listed multiple items coded as "honest" for the "good" list and 9 individuals listed two items coded as "dishonest" for the "bad" list. Again, these individuals are treated as listing each only once.

		Honesty in Good List			**Dishonesty in Bad List**		
Site	*N*	*n*	Prop.	M(SD)	*n*	Prop.	M(SD)
Coastal Tanna	44	7	0.16	0.75 (0.20)	6	0.14	0.69 (0.22)
Hadza	68	8	0.12	0.52 (0.29)	10	0.15	0.59 (0.29)
Inland Tanna	76	1	0.01	0.25* (—)	2	0.03	0.71 (0.06)
Lovu, Fiji	76	40	0.53	0.73 (0.31)	41	0.54	0.65 (0.26)
Mauritius	91	4	0.04	0.57 (0.34)	4	0.04	0.55 (0.34)
Marajó	77	16	0.21	0.65 (0.31)	11	0.14	0.51 (0.35)
Tyva Republic	79	24	0.30	0.83 (0.20)	38	0.48	0.74 (0.28)
Yasawa, Fiji	72	4	0.06	0.40 (0.16)	18	0.25	0.72 (0.26)
Total	583	104	0.18	0.70 (0.29)	130	0.22	0.67 (0.28)

Table 6: **Summary statistics for individuals per site who participated in the experiment and listed "honest" and "dishonest" in the good and bad list respectively.** *N* denotes number in sub-sample, *n* denotes those who listed "honest" and "dishonest," and Prop. refers to proportion of sub-sample listing these items. Means (M) and standard deviations (SD) are of item salience. *Note that this value reflects the salience score of the single individual.

5.1.4. Model

Again, elsewhere (Purzycki et al., 2016a,b), we found that individuals' beliefs about morally concerned deities' punishment and knowledge breadth contributed to fairer play in a wide variety of model specifications. We also found that the more children people had, the more likely they were to allocate more coins to themselves and local communities. As some participants in the present study

played in a treatment condition using various religious primes (with no overall effects detected), we hold this and game order (Local Community Game first = 1) constant in the regressions (see supplements for further details). In order to hold constant any effects for recognizing what the game was about, we created an indicator variable where values of 1 denote when participants thought the experiment was about fairness, honesty, and/or cheating ($n = 31$; 5% of the sample).

Here, we build upon the "Reduced models" in (Purzycki et al., in press). These reduced models were the result of backward-selected full models, had the lowest variance inflation factors, and largest sample sizes of any other model specification. We develop these models and their application in a few important ways. First, we incorporate moral models and culture as predictor variables at individual and group levels, respectively. Here, the group level refers to within-sample ubiquity. Second, we take group-level variation into account using varying effects. Rather than hold intercultural variation constant, we incorporate it into our modeling structure (Gelman, 2006; Nezlek, 2010; Pinheiro and Bates, 2000). Third, we formally develop statistical models (see supplements) in a Bayesian framework. Fourth, using prior defined distributions, we impute our missing data.

Our outcome variable is the binomially distributed count of allocating coins to the socially distant cup. As mentioned above, this offers the strongest test of fairer play as the chances of allocating a coin to geographically distant people better approximates to playing fairly than do allocations to self- or local community (Hruschka et al., 2014; Purzycki et al., 2016b). The log-odds of each allocation is defined as a linear combination of:

1. Varying intercepts for individual and group
2. Varying slopes by group for individual-level responses to moral models, moralistic gods' punishment, and moralistic gods' knowledge breadth
3. Fixed slopes by group for group-level average responses to moral models, moralistic gods' punishment, and moralistic gods' knowledge breadth
4. Simple effects for religious prime condition, order, game understanding check, game type, and number of children

The group-level responses are given their own statistical models, as they are not observed but rather must be inferred from the individual responses at each site. Rather than use simple fixed indices of group-level variation (e.g., the mean value for the cultural variables in each site), we infer them from the sample of individual statements. Then we simultaneously use the posterior distribution of each in the model. This retains all uncertainty so that we do not limit ourselves to point estimates that may lead to misleadingly false precision where there is actually a distribution of values within communities. For each set of beliefs—moral models, moralistic gods' punishment and knowledge breadth—we simultaneously estimate a varying intercept representing the average of each site's individual responses and use this intercept, with all associated uncertainty, as a predictor in the main model. This is analogous to a measurement error model, in which the group-level predictors are measured with error. In principle, then, our model is four simultaneous regressions: a main binomial regression predicting coin assignments and three varying intercept regressions predicting individual responses by field site. Formal details of the model are included in the supplementary materials.

With one exception, each individual played both games. We therefore restructured the data set to include two duplicate participant-by-variable matrices and included a binary variable denoting which game it represents ("Self Game" = 1). We included all values for coins to distant coreligionists in a single vector and the cups designated for the local co-ethnic and participants in

another single vector. The Hadza were not asked the game understanding check question, so we marginalized over these and imputed other missing values (see script for details, and Supplements for alternate imputation strategies).

We fit this model using the R package `rethinking` (version 1.71) (McElreath, 2016, 2017) and `rstan` version 2.17.3 (Stan Development Team, 2017). We assessed chain convergence by inspecting traceplots, $\hat{R}$ values, and the number of effective samples, and encountered no problems in sampling. The supplementary materials include further analyses in a frequentist statistical framework and analyses of item salience.

5.2. Results

5.2.1. Individual-Level Effects

Table 7 reports the main models. Values are exponentiated mean estimates (OR) and 95% credibility intervals (CI). We have highlighted the individual-level effects in Figure 2. As indicated by the intercept, individuals predictably bias allocations in favor of themselves and their local community across the eight sampled populations. As suggested by the Game variable, people favor themselves slightly more than they favor their local communities (i.e., the odds and range of the effect is trending toward values <1.00). Moreover, the effect of the number of children people have is trending towards favoritism for the players themselves and their local communities.

	OR [95% CI]	**OR [95% CI]**	**OR [95% CI]**
(Dis)honesty summation	1.03 [0.96, 1.12]	—	1.03 [0.97, 1.09]
Moralistic gods' punishment	1.09 [0.93, 1.29]	1.10 [0.95, 1.26]	1.10 [0.96, 1.25]
Moralistic gods' knowledge	1.25 [1.03, 1.53]	1.25 [1.04, 1.51]	1.23 [1.02, 1.51]
Number of children	0.88 [0.70, 1.08]	0.86 [0.70, 1.06]	0.87 [0.70, 1.08]
Condition (treatment = 1)	0.97 [0.90, 1.05]	0.97 [0.89, 1.04]	0.96 [0.89, 1.04]
Local Game played first = 1	1.02 [0.96, 1.09]	1.02 [0.96, 1.09]	1.02 [0.95, 1.10]
Game about honesty? (yes = 1)	1.00 [0.85, 1.20]	1.02 [0.87, 1.22]	1.02 [0.86, 1.20]
Game (Self Game = 1)	0.97 [0.93, 1.02]	0.97 [0.93, 1.02]	0.97 [0.94, 1.01]
Group-level moral models	1.30 [0.51, 3.59]	—	—
Group-level gods' punishment	1.65 [0.34, 6.25]	2.15 [0.59, 6.46]	2.23 [0.59, 6.55]
Group-level gods' knowledge	1.10 [0.29, 4.13]	1.18 [0.36, 4.01]	1.13 [0.38, 3.41]

Table 7: **Exponentiated mean estimates (OR) and 95% credibility intervals (CI) of chances of allocating a coin to geographically distant co-religionists**. The left-most model is the full model, the center model removes all moral variables, and the right-most model includes only individual-level moral models.

Our focal individual-level variables—moral models, deities' punishment and knowledge breadth—all contribute to increasing the chances of allocating a coin to the geographically distant players. Note here that the strongest effect is gods' attributed knowledge breadth; the more individuals claim gods know, the more likely they are to allocate to the distant cup. Gods' punishment also has an effect in the same direction, though not as obviously strong. Previous results (Purzycki et al., 2016b) indicated that punishment predicted larger allocations than knowledge. However, these previous models treated field site as a simple effect and did not allow any variables to have differential effects across sites. Moreover, they considered only complete cases. Here, we allow these factors to have differential effects across sites while estimating the effects of individual-level factors on fair behavior.

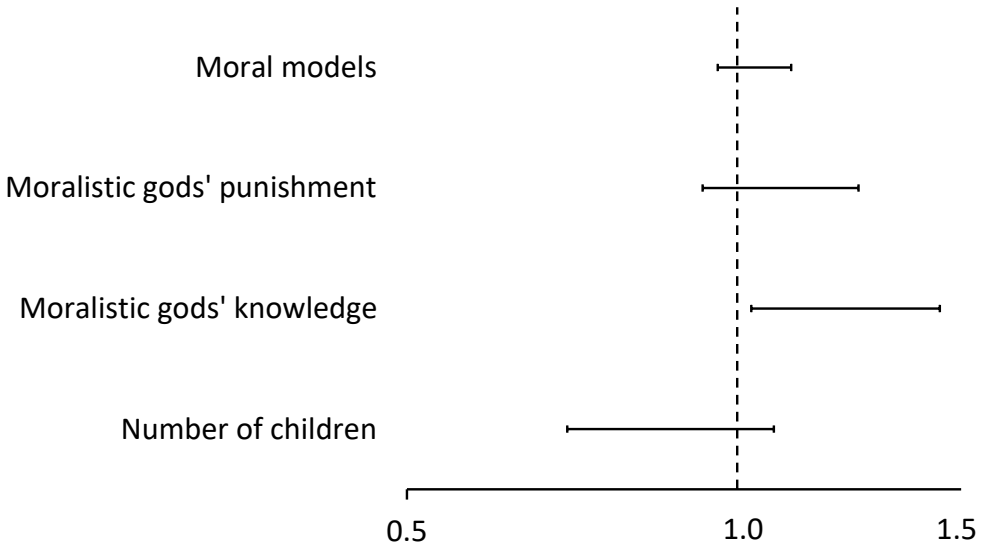

Figure 2: **Exponentiated 95% credibility intervals of mean estimates of individual-level effects of full model from table 7.** Horizontal axis is on a logarithmic scale. The dotted vertical line indicates the threshold of no effect where variables with no reliable effects would have error symmetry around 1.0. Effects to the right of 1.0 predict greater odds of allocating a coin to geographically distant players whereas effects to the left indicate decreased odds in such allocations.

The content of moral models influences game play; individuals who listed (dis)honesty are more likely to play fairly (i.e., there is a 3% greater chance of allocating a coin to distant players). Note, however, while the bulk of the probability mass is > 1, this effect is notably slight by comparison to religious beliefs. As indicated by the relatively narrower intervals, it is, however, better estimated by the model. In the supplements, among a variety of other model specifications, we show that salience of (dis)honesty has a similar relationship to behavioral outcome; individuals who list (dis)honesty earlier in lists are more likely to allocate coins to the distant players.

5.2.2. Group-Level Effects

Table 7 also includes the average contribution of moral and religious culture on allocations. Figure 3 illustrates a projection of these group-level effects, assuming participants have no children and answered all questions at the half-way mark (in this case 0.5). This includes the free-list summations, which are inverse logit transformed to put them on the same scale as the other cultural variables. Group-level beliefs in moralistic gods' punishment (OR = 1.69, 95% CI = [0.39, 6.68]) exhibits the clearest positive trend (red in Figure 3). In other words, the ubiquity of beliefs that gods punish in an individual's community positively predicts his or her fair behavior in the game (albeit with a range of error). Cultural ubiquity of (dis)honesty in moral models (green in Figure 3) and gods' knowledge breadth (blue in Figure 3) show no reliable effect on allocations (i.e., their credibility intervals are more symmetrical around 1.00).

The strongest conclusion to be drawn from the main results is that—as indicated by the wide intervals—these are poorly estimated factors. While the odds ratios (exponentiated mean estimates) appear to be high, the 95% credibility interval width is quite broad and it is difficult to conclude that culture–the prevalence of certain kinds of information within a community–has a systematic effect on individual behavior cross-culturally. Using more liberal-but-standard analyses, however,

we cautiously show that for populations where (dis)honesty is infrequently listed, models predict that increasing its cultural prevalence brings allocations to distant players in this sites to baseline, cross-site allocation levels (Supplements section 4.2.2, Figure S1). While individually held cultural information predicts individual outcomes, it remains less clear as to how cultural prevalence of that information does.

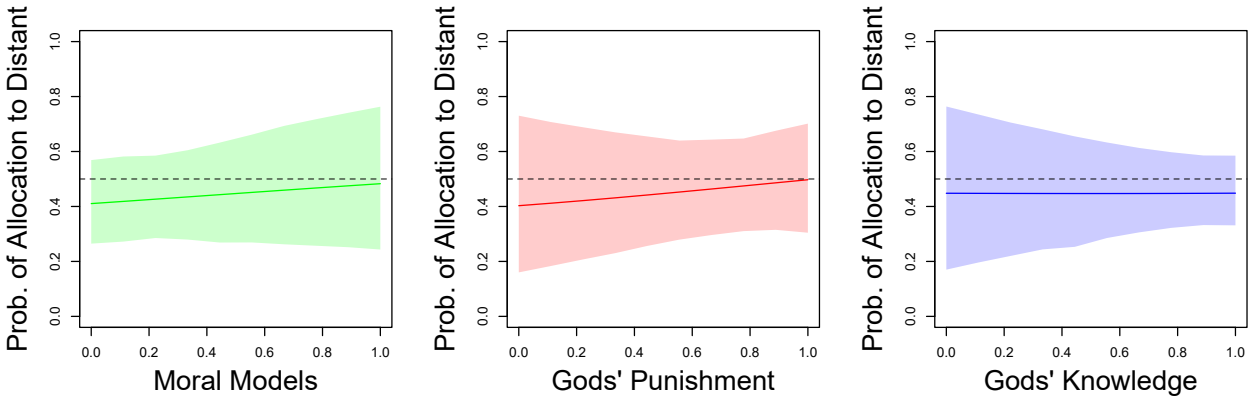

Figure 3: **Projected effects of cultural prevalence of (dis)honesty (green), gods' punishment (red), and gods' knowledge (blue) on probability of allocating a coin to distant players.** Results are from full model in Table 7. Shading is 95% percentile intervals. Reference line is at the 50% mark indicating fair play. Across projections, number of children are held at zero, and all other values are held at 0.5 (the halfway point for each variable, including the inverse logit transformed values for moral models), varying only the group level values between 0 and 1.

6. Discussion

The studies presented here give new insight into the relationship between morality and the kind of broad cooperation unique to humans. Among our diverse samples, the most salient and ubiquitous components of moral culture revolve around reciprocity, cooperation, honesty, and dishonesty. This cross-cultural ethnographic data empirically confirm that moral culture is more associated with the costs and benefits of social life (Alexander, 1987; Fessler et al., 2015; Greene, 2013; Tomasello and Vaish, 2013; Trivers, 1971) than with concerns of "justice" and "rights" (Turiel, 1983, 2006). We also found that individuals' moral models predict honest behavior towards geographically distant individuals, but their effects were not as strong as the effects of religious beliefs. As participants exhibited this behavior towards individuals they would never likely meet, our results also confirm the important role that individuals' beliefs and values have on human sociality by restraining selfish behavior. Group-level moral and religious culture, however, are *not* clearly associated with individuals' moral behavior. Below, we discuss various facets of the moral system in light of these results.

6.1. Morality in mind

While we found an individual-level effect of moral models on honest, rule-following behavior, the effect itself was quite small. This may have been due to the kind of data our methods elicit; open-ended questions require categorization for analysis, which may have introduced bias. However, field researchers collected and translated our ethnographic data which was checked multiple times for quality and consistency across research assistants. Moreover, our models exhibited considerable

precision in estimating (dis)honesty's effect on behavior. We also considered other notions beyond (dis)honesty that we might assume to be task-relevant. However, moral model components such as "cheating," "disobedience," or "fairness" were either concentrated in a few communities or rarely listed at all, thus making it difficult to reliably assess their effect on behavior in a global cross-cultural study such as ours. Using the data presented here to design scales for measuring individuals' moral models would strike a balance between universal applicability and local relevance (e.g., Boehm, 1979, 1980; Buchtel et al., 2015).

Moral models might be only as effective to the extent that an individual can implement them. Accordingly, measuring other critical individual-level factors (e.g., self-control; Blasi 1980) to predict moral behavior might be also appropriate. One indication of this in the present study is the cognitive salience of (dis)honesty. In the supplements, we show that the effects for individual-level salience of (dis)honesty are similar to those we find above; the *earlier* individuals list (dis)honesty, the greater the odds of allocating a coin to the distant participant (see supplements). In addition to accessibility, item salience may indicate how readily individuals can implement these values. This is consistent with the psychological literature that emphasizes the impact of quick moral intuitions over slow moral judgments (Baumard, 2016; Cone and Rand, 2014; Haidt, 2001; Lotz, 2015), suggesting that deeper motivational forces are at work.

Fair play had a stronger association with religious beliefs than with moral models. The effect of being watched and potentially punished by a transcendent being may have played a larger role in our expanded sociality than moral cultivation. It also raises questions as to whether or not human moralistic punishment is powerful enough to offset the costs of defecting on social expectations (cf. Dawkins, 2016; De Waal, 2013; Johnson, 2016). In other words, not only moral models but additional institutional and environmental factors are likely required to stabilize wider, more predictable cooperation. The secularization literature (e.g., Norris and Inglehart, 2012) and some of the aforementioned experimental work (e.g., Cronk, 2007; Gerkey, 2013) suggest that this is the case.

6.2. Moral behavior

Some argue that the evolutionary function of moral behavior is to maintain individuals' reputations in reciprocal interactions (Baumard, 2016; Baumard et al., 2013; Sperber and Baumard, 2012). Our result that gods' knowledge has the strongest and most reliable association with giving more coins to distant players is consistent with this view insofar as one's reputation matters in the eyes of a god. Beliefs about morally concerned deities that know about and punish people for immoral behavior might predict moral behavior more reliably because they harness–among other things–psychological systems responsible for reputation management and punishment avoidance (Johnson, 2016; Purzycki et al., 2016b). If concern of one's reputation in his or her community was strong enough to overcome both the anonymity afforded by these experiments and that target recipients were in no position to reciprocate, we should have seen a relationship between moral *culture* and honest allocations. That is, individuals should have been more likely to play honestly if it meant breaching widely held values (Baumard, 2016, p. 131, n. 13). We found no such association. It may be, however, that moral culture functions in ways not captured by such games (see below).

Ongoing concerns revolve around the ecological validity of economic experiments (Baumard and Sperber, 2010; Gervais, 2017; Gurven and Winking, 2008; Wiessner, 2009; Winking and Mizer, 2013). In small-scale societies, the anonymity afforded by these experiments is not always available. Our study takes advantage of the rarely-offered anonymity by examining whether or not participants

exploit the experimental context for their own gains. Moreover, recall that target recipients were geographically distant individuals with whom participants are unlikely to interact; we assessed whether or not cultural content can induce impartial and honest behavior in interactions we know are not happening regularly in our study sites. Other options for assessing interactions are likely to miss these rare encounters (Pisor and Gurven, 2016). As our religious belief measures showed a relatively strong association with allocation, this suggests that we cannot easily dismiss the value of using such games *in toto*.

6.3. Moral culture

Consistent with the precedent study examining cultural variation (Gurven et al., 2008), our results are inconclusive as to whether or not moral culture actually corresponds to individual behavior. As indicated by the wide credible intervals of our model estimates, group-level moral or religious commitments within a community do not reliably predict individual behavior. It may be the case that there are unmeasured, contextual factors that may be responsible for this model uncertainty. For instance, we do not know the relationship between moral culture and the threat of punishment for moral violations in each sample (Boyd and Richerson, 1992). We also do not have a reliable sense of moral culture's relationship with socioecological factors such as material security (Hruschka et al., 2014) or environmental harshness (Gelfand et al., 2011). As our sample is limited to eight field sites, more attention to group-level measures such as these in a larger sample would facilitate a more reliable assessment of these factors' relative contributions.

However, it may be the case that behaviors that correspond to moral virtues occur too context-specifically or situationally to be reliably evoked in experimental games (Fessler et al., 2015; Gerkey, 2013; North, 1991). We do not have a precise grasp of the components of individuals' moral models that become salient when they operate in different contexts (or whether or not they do). At the group level, rather than having a direct, measurable effect on our behavior, cultural ubiquity and institutions may only facilitate learning the rules and norms for successfully navigating social life (cf. Brodbeck et al., 2013; Cohn et al., 2014; Cronk, 2007; Lesorogol, 2007; Gerkey, 2013; Smaldino, 2014), thus making them more salient *in situ*. While our results suggest that the composition of mental models predicts cooperative behavior that transcends what might otherwise be parochial boundaries (e.g., Hruschka et al., 2014; Pisor and Gurven, 2016), examining *when*, and *where*, and *to whom* participants claim these moral prescriptions apply would be a logical next step for future inquiry (cf. Fessler et al., 2015).

6.4. Evolution of moral systems

Social systems have long been held to structure human interactions, but their mechanics are rarely detailed with empirical data. We assessed some components of moral systems here by focusing on individual models, local culture, and their contribution to honest behavior between members of disparate communities. Like any social system, moral systems are the aggregate output of the complex interactions between deeper cognitive adaptations and our socioecological environments (Barrett, 2014; Kiper and Sosis, 2014). Further consideration of deeper psychological systems (Cosmides and Tooby, 2005; Curry, 2016; Graham et al., 2011) and their relationship to culture and institutions (Gerkey, 2013; Stagnaro et al., 2017) is necessary to further assess the dynamism between moral systems and human cooperation.

As is true with all cross-sectional and correlational research, we cannot satisfactorily explain variation or account for the causal links and feedback between facets of moral systems. The

variability we see in moral culture may reflect variation in the challenges people collectively face together, but whether or not it contributes to the resolution of those challenges remains an open question (Alexander, 1987; Curry, 2016; Greene, 2013; Haidt and Kesebir, 2010). As is the case for most traits, uncovering the genesis of moral systems is a difficult-if not impossible-task, but longitudinal research on the topic would be better able to address the links between cognitive adaptations, culture, and environment by tracking individual- and group-level moral models and behavior, including how they change, from which sources they appear to develop, and the forces at work in their selective retention.

Acknowledgments

This work was made possible by the Cultural Evolution of Religion Research Consortium, funded by a SSHRC partnership grant and the John Templeton Foundation awarded to J.H. and A.N., and support from the Max Planck Institute for Evolutionary Anthropology (MPI). J.H. thanks CIFAR. We thank Adam Barnett, Nicholas Chan, Radek Kundt, Eva Kundtová Klocová, Tiffany Lai, Peter Manŏ, and Ted Slingerland. We also thank the Department of Human Behavior, Ecology and Culture at MPI for their feedback, particularly Bret Beheim for his help with the by-site plots in the supplements. We also thank the reviewers and Dan Fessler for their critical feedback.

Author Contributions

B.G.P. conceived, initiated, and managed this project, wrote the bulk of the manuscript, supplements, and R code, performed analysis and contributed to model development. A.P. contributed to writing the manuscript, supplements, and R code. R.M. contributed to main model development, its description, corollary R code, and group-level plots. C.A., Q.D.A., E.C., R.A.M., B.G.P., A.K.W., and D.X. collected data and contributed to protocol design with J.H. and A.N. All authors provided feedback.

Supplementary Materials

Supplementary materials including data, R scripts, and methods protocols to this article can be found online at `https://github.com/bgpurzycki/Moral-Models-Moral-Behavior` and at Purzycki et al. (2017, 2016a).

Research Ethics

This project was initially approved by the University of British Columbia's Behavioural Research Ethics Board (#H13-00671) and subsequently approved by the ethical review boards at the home university of each researcher who collected data.

Competing Interests

We declare that we have no competing interests.

References

Alexander, R. D., 1987. The Biology of Moral Systems. Aldine Transaction, New Brunswick.

Apicella, C. L., Marlowe, F. W., Fowler, J. H., Christakis, N. A., 2012. Social networks and cooperation in hunter-gatherers. Nature 481 (7382), 497.

Ashton, M. C., Lee, K., de Vries, R. E., 2014. The hexaco honesty-humility, agreeableness, and emotionality factors a review of research and theory. Personality and Social Psychology Review 18 (2), 139–152.

Barrett, H. C., 2014. The shape of thought: How mental adaptations evolve. Oxford University Press.

Barrett, H. C., Bolyanatz, A., Crittenden, A. N., Fessler, D. M., Fitzpatrick, S., Gurven, M., Henrich, J., Kanovsky, M., Kushnick, G., Pisor, A., et al., 2016. Small-scale societies exhibit fundamental variation in the role of intentions in moral judgment. Proceedings of the National Academy of Sciences 113 (17), 4688–4693.

Baumard, N., 2016. The origin of fairness. Oxford University Press, Oxford.

Baumard, N., André, J.-B., Sperber, D., 2013. A mutualistic approach to morality: The evolution of fairness by partner choice. Behavioral and Brain Sciences 36 (01), 59–78.

Baumard, N., Sperber, D., 2010. Weird people, yes, but also weird experiments. Behavioral and brain sciences 33 (2-3), 84–85.

Bernard, H. R., May 2011. Research Methods in Anthropology, 5th Edition. AltaMira Press, Lanham.

Blasi, A., 1980. Bridging moral cognition and moral action: A critical review of the literature. Psychological bulletin 88 (1), 1.

Boehm, C., Jan. 1979. Some Problems with Altruism in the Search for Moral Universals. Behavioral Science 24 (1), 15–24.

Boehm, C., 1980. Exposing the Moral Self in Montenegro: The Use of Natural Definitions to Keep Ethnography Descriptive. American Ethnologist 7 (1), 1–26.

Bowles, S., Gintis, H., 2003. Origins of human cooperation. Genetic and cultural evolution of cooperation 2003, 429–43.

Boyd, R., 2017. A different kind of animal: how culture transformed our species: how culture transformed our species. Princeton University Press.

Boyd, R., Richerson, P. J., 1988. Culture and the evolutionary process. University of Chicago Press, Chicago.

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

Boyd, R., Richerson, P. J., 2009. Culture and the evolution of human cooperation. Philosophical Transactions of the Royal Society of London B: Biological Sciences 364 (1533), 3281–3288.

Brodbeck, F. C., Kugler, K. G., Reif, J. A., Maier, M. A., 2013. Morals matter in economic games. PloS one 8 (12), e81558.

Brown, D., Jan. 1991. Human universals, 1st Edition. McGraw-Hill Humanities/Social Sciences/Languages, Boston.

Buchtel, E. E., Guan, Y., Peng, Q., Su, Y., Sang, B., Chen, S. X., Bond, M. H., 2015. Immorality east and west: Are immoral behaviors especially harmful, or especially uncivilized? Personality and Social Psychology Bulletin 41 (10), 1382–1394.

Caton, C. E., 1963. In what sense and why 'ought'–judgements are universalizable. The Philosophical Quarterly (1950-) 13 (50), 48–55.

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

Cohn, A., Fehr, E., Maréchal, M. A., 2014. Business culture and dishonesty in the banking industry. Nature 516 (7529), 86–89.

Cone, J., Rand, D. G., 2014. Time pressure increases cooperation in competitively framed social dilemmas. PLoS one 9 (12), e115756.

Cosmides, L., Tooby, J., 2005. Neurocognitive adaptations designed for social exchange. The handbook of evolutionary psychology, 584–627.

Cronk, L., 1994. Evolutionary theories of morality and the manipulative use of signals. Zygon 29 (1), 81–101.

Cronk, L., Sep. 2007. The influence of cultural framing on play in the trust game: a Maasai example. Evolution and Human Behavior 28 (5), 352–358.

Curry, O. S., 2016. Morality as cooperation: A problem-centred approach. In: Shackelford, T. K., Hansen, R. D. (Eds.), The Evolution of Morality. Springer, pp. 27–51.

Curry, O. S., Chesters, M. J., Van Lissa, C. J., n.d. Mapping morality with a compass: Testing the theory of 'morality as cooperation' with a new questionnaire. Cited with permission.
URL `https://osf.io/x7dg3/`

D'Andrade, R., 2006. Commentary on searle's social ontology: Some basic principles culture and institutions. Anthropological Theory 6 (1), 30–39.

D'Andrade, R. G., Jul. 1981. The cultural part of cognition. Cognitive Science 5 (3), 179–195.

Darwin, C., 1871. The descent of man. Penguin Classics, Princeton, N.J.

Dawkins, R., 2016. The god delusion. Random House, New York.

De Waal, F., 2013. The bonobo and the atheist. W.W. Norton and Company, New York.

Ensminger, J., Henrich, J., 2014. Experimenting with social norms: Fairness and punishment in cross-cultural perspective. Russell Sage Foundation.

Fehr, E., Fischbacher, U., Gächter, S., 2002. Strong reciprocity, human cooperation, and the enforcement of social norms. Human Nature 13 (1), 1–25.

Fessler, D. M. T., Barrett, H. C., Kanovsky, M., Stich, S., Holbrook, C., Henrich, J., Bolyanatz, A. H., Gervais, M. M., Gurven, M., Kushnick, G., Pisor, A. C., Rueden, C. v., Laurence, S., Aug. 2015. Moral parochialism and contextual contingency across seven societies. Proc. R. Soc. B 282 (1813), 20150907.

Fischbacher, U., Föllmi-Heusi, F., 2013. Lies in disguisean experimental study on cheating. Journal of the European Economic Association 11 (3), 525–547.

Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., Duan, L., Almaliach, A., Ang, S., Arnadottir, J., et al., 2011. Differences between tight and loose cultures: A 33-nation study. science 332 (6033), 1100–1104.

Gelman, A., 2006. Multilevel (hierarchical) modeling: What it can and cannot do. Technometrics 48 (3), 432–435.

Gerkey, D., 2013. Cooperation in Context: Public Goods Games and Post-Soviet Collectives in Kamchatka, Russia. Current Anthropology 54 (2), 144–176.

Gervais, M. M., 2017. Rich economic games for networked relationships and communities: Development and preliminary validation in yasawa, fiji. Field Methods 29 (2), 113–129.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., Ditto, P. H., 2013. Moral Foundations Theory: The Pragmatic Validity of Moral Pluralism. Advances in Experimental Social Psychology 47, 55–130.

Graham, J., Meindl, P., Koleva, S., Iyer, R., Johnson, K. M., 2015. When values and behavior conflict: Moral pluralism and intrapersonal moral hypocrisy. Social and Personality Psychology Compass 9 (3), 158–170.

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., Ditto, P. H., 2011. Mapping the moral domain. Journal of personality and social psychology 101 (2), 366.

Greene, J., Oct. 2013. Moral Tribes: Emotion, Reason, and the Gap Between Us and Them. Penguin, New York.

Gurven, M., Von Rueden, C., Massenkoff, M., Kaplan, H., Lero Vie, M., 2013. How universal is the big five? testing the five-factor model of personality variation among forager–farmers in the bolivian amazon. Journal of personality and social psychology 104 (2), 354.

Gurven, M., Winking, J., 2008. Collective action in action: Prosocial behavior in and out of the laboratory. American Anthropologist 110 (2), 179–190.

Gurven, M., Zanolini, A., Schniter, E., 2008. Culture sometimes matters: Intra-cultural variation in pro-social behavior among tsimane amerindians. Journal of Economic Behavior & Organization 67 (3-4), 587–607.

Haidt, J., 2001. The emotional dog and its rational tail: a social intuitionist approach to moral judgment. Psychological review 108 (4), 814.

Haidt, J., Kesebir, S., 2010. Morality. In: Fiske, S., Gilbert, D. (Eds.), Handbook of Social Psychology, 5th Edition. Wiley, Hoboken, pp. 797–832.

Handwerker, W. P., 2001. Quick ethnography. Altamira Press, Lanham.

Henrich, J., 2000. Does culture matter in economic behavior? ultimatum game bargaining among the machiguenga of the peruvian amazon. The American Economic Review 90 (4), 973–979.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H. (Eds.), 2004. Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies. Oxford University Press, New York.

Henrich, J., Henrich, N., 2014. Fairness without punishment: Behavioral experiments in the yasawa islands, fiji. In: Ensminger, J., Henrich, J. (Eds.), Experimenting with social norms: Fairness and punishment in cross-cultural perspective. SAGE, pp. 225–258.

Hofmann, W., Wisneski, D. C., Brandt, M. J., Skitka, L. J., 2014. Morality in everyday life. Science 345 (6202), 1340–1343.

Hruschka, D., Efferson, C., Jiang, T., Falletta-Cowden, A., Sigurdsson, S., McNamara, R., Sands, M., Munira, S., Slingerland, E., Henrich, J., Dec. 2014. Impartial institutions, pathogen stress and the expanding social network. Human Nature 25 (4), 567–579.

Jamieson-Lane, A., Purzycki, B. G., 2016. AnthroTools: Some custom tools for anthropology. R package version 0.8.

Jiang, T., 2013. Cheating in mind games: The subtlety of rules matters. Journal of Economic Behavior & Organization 93, 328–336.

Johnson, D., 2016. God is watching you: How the fear of god makes us human. Oxford University Press, USA.

Kant, I., 1997 [1785]. Groundwork of the Metaphysics of Morals. Cambridge University Press, Cambridge.

Kiper, J., Sosis, R., 2014. Moral intuitions and the religious system: An adaptationist account. Philosophy, Theology and the Sciences 1 (2), 172–199.

Kurzban, R., Dukes, A., Weeden, J., 2010. Sex, drugs and moral goals: Reproductive strategies and views about recreational drugs. Proceedings of the Royal Society of London B: Biological Sciences 277 (1699), 3501–3508.

Lesorogol, C. K., 2007. Bringing norms in: The role of context in experimental dictator games. Current Anthropology 48 (6), 920–926.

Lotz, S., 2015. Spontaneous giving under structural inequality: Intuition promotes cooperation in asymmetric social dilemmas. PLoS One 10 (7), e0131562.

Machery, E., Mallon, R., 2010. Evolution of morality. The moral psychology handbook, 3.

McElreath, R., 2016. Statistical Rethinking: A Bayesian Course with Examples in R and Stan. CRC Press, New York.

McElreath, R., 2017. rethinking: Statistical Rethinking book package. R package version 1.63.

McNamara, R. A., Norenzayan, A., Henrich, J., Jan. 2016. Supernatural punishment, in-group biases, and material insecurity: experiments and ethnography from Yasawa, Fiji. Religion, Brain & Behavior 6 (1), 34–55.

Mizzoni, J., 2009. 'the social instincts naturally lead to the golden rule': The ethics of charles darwin. Teorema: Revista Internacional de Filosofía, 123–133.

Nezlek, J. B., 2010. Multilevel modeling and cross-cultural research. Cross-Cultural research methods in psychology. Oxford.

Norris, P., Inglehart, R., 2012. Sacred and Secular: Religion and Politics Worldwide, 2nd Edition. Cambridge University Press, Cambridge.

North, D. C., 1991. Institutions. Journal of Economic Perspectives 5 (1), 97–112.

Perry, S. L., 2017. Not practicing what you preach: Religion and incongruence between pornography beliefs and usage. The Journal of Sex Research, 1–12.

Pinheiro, J. C., Bates, D. M., 2000. Linear mixed-effects models: basic concepts and examples. Mixed-effects models in S and S-Plus, 3–56.

Pisor, A. C., Gurven, M., 2016. Risk buffering and resource access shape valuation of out-group strangers. Scientific Reports 6.

Purzycki, B., Apicella, C., Atkinson, Q., Cohen, E., Henrich, J., McNamara, R., Norenzayan, A., Willard, A., Xygalatas, D., 2017. Evolution of religion and morality project dataset (wave 1). URL `http://dx.doi.org/10.7910/DVN/RT5JTV`

Purzycki, B. G., 2011. Tyvan *cher eezi* and the socioecological constraints of supernatural agents' minds. Religion, Brain & Behavior 1 (1), 31–45.

Purzycki, B. G., Apr. 2016. The Evolution of Gods' Minds in the Tyva Republic. Current Anthropology 57 (S13), S88–S104.

Purzycki, B. G., Apicella, C., Atkinson, Q. D., Cohen, E., McNamara, R. A., Willard, A. K., Xygalatas, D., Norenzayan, A., Henrich, J., 2016a. Cross-cultural dataset for the evolution of religion and morality project. Scientific Data 3, 160099.

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

Purzycki, B. G., Jamieson-Lane, A., 2016. Anthrotools an r package for cross-cultural ethnographic data analysis. Cross-Cultural Research 51 (1), 51–74.

Purzycki, B. G., Norenzayan, A., Apicella, C., Atkinson, Q. D., Cohen, E., Henrich, J., McNamara, R. A., Willard, A. K., Xygalatas, D., in press. The evolution of religion and morality: A synthesis of ethnographic and experimental evidence from eight societies. Religion, Brain and Behavior.

R Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Richerson, P., Baldini, R., Bell, A. V., Demps, K., Frost, K., Hillis, V., Mathew, S., Newton, E. K., Naar, N., Newson, L., et al., 2016. Cultural group selection plays an essential role in explaining human cooperation: A sketch of the evidence. Behavioral and Brain Sciences 39, e30.

Roth, A. E., Prasnikar, V., Okuno-Fujiwara, M., Zamir, S., 1991. Bargaining and market behavior in jerusalem, ljubljana, pittsburgh, and tokyo: An experimental study. The American Economic Review, 1068–1095.

Rozin, P., 1999. The process of moralization. Psychological Science 10 (3), 218–221.

Schwartz, S. H., Nov. 2007. Universalism Values and the Inclusiveness of Our Moral Universe. Journal of Cross-Cultural Psychology 38 (6), 711–728.

Searle, J. R., 1995. The construction of social reality. Simon and Schuster.

Shweder, R. A., Much, N. C., Mahapatra, M., Park, L., 1997. The "big three" of morality (autonomy, community, divinity) and the "big three" explanations of suffering. In: Brandt, A. M., Rozin, P. (Eds.), Morality and health. Taylor & Frances/Routledge, pp. 119–169.

Smaldino, P. E., 2014. The cultural evolution of emergent group-level traits. Behavioral and Brain Sciences 37 (3), 243–254.

Smith, C. E., Blake, P. R., Harris, P. L., 2013. I should but i wont: Why young children endorse norms of fair sharing but do not follow them. PloS one 8 (3), e59510.

Smith, J. J., 1993. Using anthopac 3.5 and a spreadsheet to compute a free-list salience index. Cultural Anthropology Methods 5 (3), 1–3.

Smith, J. J., Borgatti, S. P., Dec. 1997. Salience Counts—And So Does Accuracy: Correcting and Updating a Measure for Free-List-Item Salience. Journal of Linguistic Anthropology 7 (2), 208–209.

Smith, J. J., Furbee, L., Maynard, K., Quick, S., Ross, L., 1995. Salience Counts: A Domain Analysis of English Color Terms. Journal of Linguistic Anthropology 5 (2), 203–216.

Smith, K. D., Smith, S. T., Christopher, J. C., 2007. What defines the good person? cross-cultural comparisons of experts' models with lay prototypes. Journal of Cross-Cultural Psychology 38 (3), 333–360.

Sperber, D., 1996. Explaining culture: A naturalistic approach. Blackwell Oxford.

Sperber, D., Baumard, N., 2012. Moral reputation: An evolutionary and cognitive perspective. Mind & Language 27 (5), 495–518.

Sripada, C. S., Stich, S., 2006. A framework for the psychology of norms. In: Carruthers, P., Laurence, S., Stich, S. (Eds.), The Innate Mind: Culture and Cognition. Vol. 2. Oxford University Press, Oxford, pp. 280–301.

Stagnaro, M. N., Arechar, A. A., Rand, D. G., 2017. From good institutions to generous citizens: Top-down incentives to cooperate promote subsequent prosociality but not norm enforcement. Cognition.

Stan Development Team, 2017. RStan: the R interface to Stan. R package version 2.16.2. URL `http://mc-stan.org/`

Tomasello, M., Vaish, A., 2013. Origins of human cooperation and morality. Annual review of psychology 64, 231–255.

Trivers, R. L., 1971. The evolution of reciprocal altruism. The Quarterly review of biology 46 (1), 35–57.

Turiel, E., 1983. The Development of Social Knowledge: Morality and Convention. Cambridge University Press.

Turiel, E., 2006. Thought, emotions, and social interactional processes in moral development. In: Killen, M., Smetana, J. (Eds.), Handbook of moral development. Lawrence Erlbaum Associates, Publishers, pp. 7–35.

Wierzbicka, A., 1994. Semantic primitives across languages: A critical review. Semantic and lexical universals: Theory and empirical findings, 445–500.

Wiessner, P., 2009. Experimental games and games of life among the Ju/'hoan bushmen. Current Anthropology 50 (1), 133–138.

Winking, J., Mizer, N., 2013. Natural-field dictator game shows no altruistic giving. Evolution and Human Behavior 34 (4), 288–293.

Supplementary Materials for: The Cognitive and Cultural Foundations of Moral Behavior

Benjamin Grant Purzycki[a,*], Anne C. Pisor[a], Coren Apicella[b], Quentin Atkinson[c,d], Emma Cohen[e,f], Joseph Henrich[g], Richard McElreath[a], Rita A. McNamara[h], Ara Norenzayan[i], Aiyana K. Willard[e], Dimitris Xygalatas[j]

[a]*Department of Human Behavior, Ecology and Culture, Max Planck Institute for Evolutionary Anthropology, DE*
[b]*Department of Psychology, University of Pennsylvania, USA*
[c]*Department of Psychology, University of Auckland, NZ*
[d]*Max Planck Institute for the Science of Human History, DE*
[e]*Institute of Cognitive and Evolutionary Anthropology, University of Oxford, UK*
[f]*Wadham College, University of Oxford, UK*
[g]*Department of Human Evolutionary Biology, Harvard University, USA*
[h]*School of Psychology, Victoria University of Wellington, NZ*
[i]*Department of Psychology, University of British Columbia, CA*
[j]*Department of Anthropology, University of Connecticut, USA*

Contents

*Corresponding author
Email address: benjamin_purzycki@eva.mpg.de (Benjamin Grant Purzycki)

1. Protocols, Data, and R Code

The methodological protocols, data sets, and R code are publicly available at: `https://github.com/bgpurzycki/Moral-Models-Moral-Behavior`. The R code includes step-by-step instructions on how to analyze and restructure the free-list data according to the appropriate analyses including: 1) calculations of item salience for free-list data; 2) assessing universal and local models of good and bad; 3) extracting binary values for listing "honesty" and "dishonesty"; 4) adding the salience and binary values to the main data set; and 5) code for regressing these and other values on coin allocations for both main and supplementary analyses.

2. Sample

Below, we briefly detail the populations with whom we conducted our work. Details for the present study are in table S1. Further details of these sites and our sampling techniques can be found at [15, 18].

Site/Sample	Moralistic Deity	Distant Player	Language of Study
Coastal Tanna	Christian god	from another Christian village	Bislama
Hadza	*Haine*	another Hadza from another camp	Hadzane/Swahili
Inland Tanna	*Kalpapen*	from another *Kastom* village	Navhaal
Lovu, Fiji	Shiva	Hindu on Vanua Levu	Fiji-Hindi/English
Mauritians	Shiva	Hindu from La Gaulette	Mauritian Creole
Marajó, Brazil	Christian god	Christian from Rondon	Portuguese
Tyva Republic	*Buddha-Burgan*	Buddhist from Ak Dovurak	Tyvan
Yasawa, Fiji	Christian god	Fijian Christians from another island	Bauan Fijian

Table S1: **Further descriptive features of each field site.**

2.1. Coastal and Inland Tanna, Vanuatu

The inhabitants of Tanna Island in Vanuatu are traditionally swidden horticulturalists although a market-based economy plays an increasingly important role on the island [2, 4]. Religious beliefs are a mix of Christianity and the traditional "Kastom" pantheon, as well as millenarian "cargo cults." The study was conducted at two sites on Tanna: a cluster of three inland, predominantly Kastom hamlets that rely almost exclusively on subsistence farming for food production, and a wealthier coastal, Christian village in which home production accounts for about two thirds of food consumption. For the Inland Tanna site, the "moralistic god" used for the survey questions was the *Kastom* creator god and culture hero, *Kalpapen* while the Coastal site used the Christian god.

2.2. Hadza, Tanzania

Living in the savannah woodlands of western Tanzania, the Hadza are a population of hunter-gatherers who largely subsist on wild game, fruits, tubers, and honey [1, 11]. While the Hadza have been described as having a minimalist form of religion, this appears to be changing. Hadza camps exhibit fission-fusion organization and camp membership is quite fluid, with individuals moving frequently between camps. Labor is divided between the sexes; men hunt and extract honey while women typically focus on gathering plants.

2.3. Lovu, Fiji

On the south Pacific island of Vanua Levu, in main island in the Fijian archipelago, the Indo-Fijians are a diaspora population brought to Fiji from India by the British as indentured workers [9, 22]. They are primarily wage laborers or sugar cane farmers. The Indo-Fijians are mostly Hindus and Muslims with a minority of Sikhs and Christians. The present sample consisted of Hindu Indo-Fijians from Lovu village on the island of Viti Levu. The participants were all wage laborers or unemployed. Though there are many gods in the Hindu tradition, the participants believed that all gods are aspects of one single God (*Bhagwan*).

2.4. Marajó, Brazil

Pesqueiro is a small fishing village on the east side of Marajó Island at the mouth of the Amazon River [6]. Subsistence is primarily market-based, relying on fish sales in the nearby town of Soure and a growing tourism industry. The majority of inhabitants identify as Catholic, though there is a minority of Evangelical Protestants.

2.5. Pointe aux Piments, Mauritius

Mauritius is a cluster of islands about 1,200 miles off the coast of southeastern Africa [5, 23]. Though historically dependent largely on sugar exports, Mauritius has developed into a diversified, market-based, monetized economy in recent decades. The main employment sectors include manufacturing, tourism, financial services, information technology, fish processing, and construction. Rural areas continue to rely on horticulture and fishing for subsistence. The study was conducted in the coastal rural village of Pointe aux Piments, which lacks industrial development. The majority of the local population is of low or middle income, employed mainly in fishing, agriculture, tourism, and other services. The village has a religiously mixed population, with Christians and Hindus each making up approximately 45% of the total. Our sample for this study consisted of Hindus.

2.6. Kyzyl, Tyva Republic (Russia)

Part of the Russian Federation, the Tyva Republic lies in southern Siberia, in the center of Asia [17, 21]. Urban Tyvans subsist primarily on a market-based economy while rural Tyvans rely significantly more on produce provided by livestock (sheep, goats, cattle, yaks). The study was conducted in the capital city of Kyzyl primarily among urban Tyvans, though some were farmers. While the majority of Tyvans identifies as Buddhist, traditional religious practices associated with shamanism, animism, and totemism have a strong presence as well.

2.7. Yasawa, Fiji

Yasawa Island lies at the northwestern corner of the Fijian archipelago [8, 12]. Yasawans subsist primarily as fisher-horticulturalists. The present sample consists mainly of villagers living closest to the only resort on Yasawa Island as of July 2013. As such, this village has had the highest population of residents with the most extensive and frequent interaction with a resort. All Yasawans identify as Christian, with a majority practicing as Wesleyan Methodists and a large minority practicing as Assemblies of God evangelicals. Additionally, their Christian beliefs and practices coexist alongside beliefs about traditional deified ancestor spirits that can bring illness, misfortune, and death to those who deviate from proper traditional Fijian social norms, often at the behest of sorcerers.

3. Moral Models

3.1. Methods

For reasons expressed in the main article (e.g., difficulty with scale designs), the most straightforward and universally applicable way to ask participants about moral models was asking them to freely list things that they think are "good" and "bad." Our protocols instructed field researchers to ask:

- *Please list up to 5 behaviors that make someone a good/virtuous/moral person.*
- *Please list up to 5 behaviors that make a bad/immoral person.*

Because of local variation and the lack of lexical equivalents to "moral" and "immoral," we required some flexibility in design. Per site English translations of the good/bad questions are as follows with language of study in parentheses. Participants were asked to "Please list up to 5 behaviors that..."

- **Hadza (Hadzane and Swahili):** "...make someone with good/bad character or behavior."

- **Coastal Tanna (Bislama):** "...make someone good/not good."
- **Inland Tanna (Nivhaal):** "...make someone good/bad."
- **Lovu, Fiji (Fiji Hindi and English):** "...make a good, virtuous, moral person/bad, immoral person."
- **Marajó, Brazil (Portuguese):** "...make a good or virtuous person/bad or immoral person."
- **Mauritius (Mauritian Creole):** "...a good, moral, and virtuous person/bad or immoral person has."
- **Tyva Republic (Tyvan):** "...make a good/bad person."
- **Yasawa, Fiji (Bauan Fijian):** "...make a good/bad person."

3.2. Codebook for free-list data

The data set (the `CERC_Moral_Model_FL.csv` file) contains 15 variables. To provide a variety of ways to code the data and as a quality check, Purzycki and two research assistants (NC and TL henceforth) independently cleaned and coded the data. We strove to ensure that items with subtle lexical differences were the same (e.g., "thievery" and "theft" would be recoded as "theft"). Purzycki encouraged assistants to cluster items conservatively and to keep their own codes consistent across sites. One small chunk of this data set was coded in error and some individuals' data were recovered after the initial coding regime (see data set and codebook). We therefore use Purzycki's recoded data in the present analyses.

- `Culture`: Name of field site
- `CERCID`: ID number of participant
- `Order`: Order in which item was listed (1 = first listed)
- `GOOD_ORIG_NC`: Original in-English translated data for "good" data used by coder NC
- `GOOD_ORIG_TL`: Original in-English translated data for "good" data used by coder TL
- `MATCH_GOOD`: Match check of original items in "good" list. "Corrected" indicates where Purzycki found inconsistencies due to a sorting error, and subsequently corrected them. "Match" indicates consistency across coders. "NewCode" indicates data that were recovered after initial coding by NC and TL.
- `GOOD_SPEC_NC`: Specific coding scheme for "good" list by coder NC (unless `MATCH_GOOD` is coded as "NewCode," see below)
- `GOOD_SPEC_TL`: Specific coding scheme for "good" list by coder TL
- `GOOD_SPEC_BP`: Specific coding scheme for "good" list by Purzycki (used in present analyses)
- `BAD_ORIG_NC`: Original in-English translated data for "bad" data used by coder NC (unless `MATCH_BAD` is coded as "NewCode," see below)
- `BAD_ORIG_TL`: Original in-English translated data for "bad" data used by coder TL
- `MATCH_BAD`: Match check of original items in "bad" list. "Corrected" indicates where Purzycki found inconsistencies due to a sorting error, and subsequently corrected them. "Match" indicates consistency across coders. "NewCode" indicates data that were recovered after initial coding by NC and TL.
- `BAD_SPEC_NC`: Specific coding scheme for "bad" list by coder NC
- `BAD_SPEC_TL`: Specific coding scheme for "bad" list by coder TL
- `BAD_SPEC_BP`: Specific coding scheme for "bad" list by Purzycki (used in present analyses)

3.3. Results

Tables S2 and S3 detail item types listed cross-culturally for the "good" and "bad" lists respectively (see Tables 3 and 4 in the main text for truncated versions). In these tables, we report only those items where Smith's S ≥ 0.10 for the sake of viewing ease. The R script includes code to view and reproduce full tables with all items where Smith's S is <0.10. Smith's S scores were calculated using the following sample sizes (good/bad lists): (1) Coastal Tanna-44/44; (2) Hadza-69/69; (3) Inland Tanna-74/78; (4) Lovu Fiji-79/79; (5) Mauritius-82/84; (6) Marajó-77/76; (7) Tyva Republic-115/117; (8) Yasawa Fiji-103/103.

Table S2: **Cross-cultural models for what makes a "good" person.** Mean (M) and sum (S) of salience scores. Only Smith's S ≥ 0.10 are reported. Recall that the denominator for M Salience is the number of Items listed, whereas the denominator in Smith's S is sample size. *Indicative of responses that include "good people" or "character" in their response.

Culture	Item	M Salience	S Salience	Smith's S
Coastal Tanna	Generous/Shares	0.74	17.67	0.40
Coastal Tanna	Respectful	0.61	13.35	0.30
Coastal Tanna	Helpful	0.71	12.00	0.27
Coastal Tanna	Loving	0.65	9.82	0.22
Coastal Tanna	Kind	0.77	6.93	0.16
Coastal Tanna	Obedient	0.56	6.75	0.15
Coastal Tanna	Hospitable	0.49	5.35	0.12
Coastal Tanna	Honest	0.75	5.25	0.12
Coastal Tanna	No Stealing	0.58	5.20	0.12
Hadza	Good*	0.84	32.60	0.47
Hadza	Loving	0.62	27.48	0.40
Hadza	Generous/Shares	0.56	23.32	0.34
Hadza	Peaceful	0.60	13.27	0.19
Hadza	Hospitable	0.64	10.25	0.15
Hadza	Intelligent	0.65	7.83	0.11
Hadza	Respectful	0.58	7.53	0.11
Inland Tanna	Hospitable	0.74	33.08	0.45
Inland Tanna	Generous/Shares	0.69	31.97	0.43
Inland Tanna	Kind	0.87	26.17	0.35
Inland Tanna	Respectful	0.49	23.98	0.32
Inland Tanna	No Stealing	0.66	18.52	0.25
Inland Tanna	Obedient	0.52	14.63	0.20
Inland Tanna	Loving	0.81	10.47	0.14
Inland Tanna	Helpful	0.41	9.38	0.13
Inland Tanna	No Swearing	0.49	8.88	0.12
Lovu Fiji	Honest	0.73	30.02	0.38
Lovu Fiji	Helpful	0.58	16.88	0.21
Lovu Fiji	Good*	0.74	15.58	0.20
Lovu Fiji	Religious Faith	0.78	10.85	0.14
Lovu Fiji	Speaking Well	0.81	8.07	0.10
Mauritius	Speaking Well	0.84	26.87	0.33
Mauritius	Helpful	0.62	24.00	0.29
Mauritius	Good*	0.71	15.70	0.19
Mauritius	Generous/Shares	0.72	15.12	0.18
Mauritius	Manners	0.73	10.90	0.13
Mauritius	Hard-working	0.65	9.80	0.12
Mauritius	Respectful	0.60	8.45	0.10
Pesqueiro	Helpful	0.70	14.07	0.18
Pesqueiro	Caring	0.68	13.65	0.18
Pesqueiro	Generous/Shares	0.65	11.03	0.14
Pesqueiro	Honest	0.65	10.47	0.14
Pesqueiro	Happy	0.69	9.63	0.13
Pesqueiro	Manners	0.77	9.20	0.12
Pesqueiro	Good*	0.68	8.87	0.12
Pesqueiro	Friendly	0.47	8.43	0.11
Pesqueiro	Humble	0.58	8.15	0.11
Tyva Republic	Good*	0.72	43.42	0.38

Continued on next page

Table S2 – *"Good" models continued from previous page*

Culture	Item	M Salience	S Salience	Smith's S
Tyva Republic	Honest	0.74	31.95	0.28
Tyva Republic	Helpful	0.61	26.25	0.23
Tyva Republic	Hard-working	0.58	24.90	0.22
Tyva Republic	Respectful	0.63	15.22	0.13
Tyva Republic	Kind	0.64	12.80	0.11
Tyva Republic	Intelligent	0.53	11.60	0.10
Tyva Republic	Humane	0.66	11.20	0.10
Yasawa Fiji	Church	0.68	40.10	0.39
Yasawa Fiji	Attitude	0.80	25.60	0.25
Yasawa Fiji	Obedient	0.64	21.60	0.21
Yasawa Fiji	Merciful	0.64	20.60	0.20
Yasawa Fiji	Good*	0.64	20.40	0.20
Yasawa Fiji	Listens	0.81	18.60	0.18
Yasawa Fiji	Hard-working	0.53	18.40	0.18
Yasawa Fiji	Generous/Shares	0.54	10.20	0.10

Table S3: **Cross-cultural models for what makes a "bad" person.** Mean (M) and sum (S) of salience scores. Only Smith's S $\geq$ 0.10 are reported. Recall that the denominator for M Salience is the number of Items listed, whereas the denominator in Smith's S is sample size. *Indicative of responses that include "bad people" or "character" in their response.

Culture	Item	M Salience	S Salience	Smith's S
Coastal Tanna	Theft	0.76	20.40	0.46
Coastal Tanna	Disrespectful	0.64	14.15	0.32
Coastal Tanna	Violent	0.66	12.57	0.29
Coastal Tanna	Disobedient	0.64	8.27	0.19
Coastal Tanna	Selfish	0.74	6.70	0.15
Coastal Tanna	Bad Language (Swearing)	0.47	5.70	0.13
Coastal Tanna	Drugs/Alcohol/Substance	0.69	4.80	0.11
Hadza	Murder	0.52	23.60	0.34
Hadza	Greedy	0.77	22.45	0.33
Hadza	Theft	0.60	21.63	0.31
Hadza	Violent	0.65	21.35	0.31
Hadza	Bad*	0.81	17.87	0.26
Hadza	Angry	0.66	6.63	0.10
Hadza	Troublemaker	0.82	6.60	0.10
Inland Tanna	Theft	0.76	28.72	0.37
Inland Tanna	Inhospitable	0.78	22.72	0.29
Inland Tanna	Disrespectful	0.47	21.20	0.27
Inland Tanna	Bad Language (Swearing)	0.63	18.25	0.23
Inland Tanna	Disobedient	0.59	15.88	0.20
Inland Tanna	Unkind	0.75	13.58	0.17
Inland Tanna	Violent	0.73	12.35	0.16
Inland Tanna	Not Feeding Others	0.64	8.35	0.11
Inland Tanna	Unhelpful	0.47	7.58	0.10
Lovu Fiji	Theft	0.82	28.73	0.36
Lovu Fiji	Dishonest	0.64	26.97	0.34
Lovu Fiji	Jealous	0.68	16.42	0.21
Lovu Fiji	Violent	0.70	14.03	0.18
Lovu Fiji	Drugs/Alcohol/Substance	0.81	8.92	0.11
Lovu Fiji	Betrayal/Backbiting	0.62	8.70	0.11
Lovu Fiji	Bad Company/Peers	0.84	8.42	0.11
Mauritius	Violent	0.65	25.28	0.30
Mauritius	Drugs/Alcohol/Substance	0.70	25.07	0.30
Mauritius	Theft	0.62	14.90	0.18
Mauritius	Doesn't Speak Well	0.70	13.98	0.17
Mauritius	Disrespectful	0.65	10.98	0.13
Mauritius	Selfish	0.64	10.85	0.13
Mauritius	Bad*	0.60	9.53	0.11
Pesqueiro	Ignorance/Stupidity/Uncultured	0.69	10.30	0.14
Pesqueiro	Bad*	0.77	10.03	0.13
Pesqueiro	Selfish	0.55	9.93	0.13
Pesqueiro	Envious	0.74	8.85	0.12
Pesqueiro	Unhelpful	0.85	8.47	0.11
Tyva Republic	Dishonest	0.70	41.25	0.35
Tyva Republic	Bad*	0.75	20.90	0.18
Tyva Republic	Drugs/Alcohol/Substance	0.63	19.02	0.16
Tyva Republic	Cruel	0.56	15.65	0.13
Tyva Republic	Envious	0.71	15.55	0.13

Continued on next page

Table S3 – *"Bad" models continued from previous page*

Culture	Item	M Salience	S Salience	Smith's S
Tyva Republic	Lazy	0.60	15.50	0.13
Tyva Republic	Theft	0.57	15.40	0.13
Yasawa Fiji	Theft	0.80	33.80	0.33
Yasawa Fiji	Disobedient	0.64	22.95	0.22
Yasawa Fiji	No Church	0.53	18.60	0.18
Yasawa Fiji	Dishonest	0.69	15.13	0.15
Yasawa Fiji	Doesn't Listen	0.64	14.07	0.14
Yasawa Fiji	Bad Language (Swearing)	0.67	12.80	0.12
Yasawa Fiji	Proud	0.66	11.25	0.11
Yasawa Fiji	Lazy	0.61	11.00	0.11
Yasawa Fiji	Drugs/Alcohol/Substance	0.79	11.00	0.11
Yasawa Fiji	Greedy	0.62	10.00	0.10

3.4. Notes on the experiment's effects on free-lists

The fact that we interviewed participants after they played games may raise suspicions that the experiments somehow influenced the content of moral models. There are a few reasons why participants were probably not more inclined to say "honest" and "dishonest" after playing a game measuring (dis)honest behavior. First, participants answered demographic questions before answering the free-list questions. This process took about 10-15 minutes. Second, the game check variable we hold constant should account for–at the very least–explicit recognition of the game's purpose. A logistic regression without considering the game check variable shows that the probability of listing "honesty" is 18% (logistic transform of -1.52). The mean estimate for the game check question was 0.62. Logistic transforming this summation yields a 29% probability of listing "honesty." Adding the mean estimate for the game check question (0.62) therefore suggests an increase of 11% (29% - 18% = 11%). Note, however, that only 9 individuals who listed "honesty" recognized that the game was about it (397 did not list honesty *and* did not indicate they understood what the game was measuring, and 22 recognized the purpose of the game, but did not list honesty). Third–as per Table 5 in the main text–there are a few sites where few individuals listed (dis)honesty, suggesting that if the game was priming the free-listing of (dis)honesty, it was not occurring systematically. Fourth, previous research [19] assessing what it means to be good shows that across seven societies, "honest" is highly salient in most contexts.

Fifth, a previous case study [13] using a similar free-list method with a different sample who didn't play the game (in the Tyva Republic) suggests there is some consistency in content across these studies. There, items were coded independently of the present project, and participants were encouraged to list 10-15 items of what makes specifically a good or bad *Tyvan* person, rather than the capped-at-5 general description of (im)moral people as was the case here. Moreover, for the present study, to ensure that recipients and players were of the same religion, Purzycki recruited participants who self-identified as Buddhist. In the previous study, recruitment was open to any ethnic Tyvan. In the previous study, the eight-most salient items for what it means to be a "good Tyvan" were: hard-working (S = 0.40), helpful (0.30), kind (0.29), modest (0.28), respectful (0.26), honest (0.22), intelligent (0.22) and having love for family (0.19). "Bad Tyvans" were: untrustworthy (S = 0.66), alcohol abusers (0.53), lazy (0.34), envious (0.20), greedy (0.19), disrespectful (0.18), cruel (0.17), and ignorant (0.14). In terms of content, both are quite similar to the present results (see Table S2). In the present Tyvan sample, "honest" had a Smith's S of 0.28 (vs. 0.22 in the previous study), a negligible difference. However, "dishonest" had an S of 0.35 (vs. 0.66 for "untrustworthy"), much lower than the previous study. This would suggest that at least in the case of Tyvans, the game could potentially reduce the salience of items, at least in the "bad" subdomain. However, "honesty" (after ignoring "good" in the good free-list task) and "dishonesty" were the *most* salient items listed for the present sample. In summary, despite methodological, coding, and sampling differences, even if the game reduced the salience of dishonesty for Tyvans, these items still had the highest salience across both studies.

4. Moral Behavior

4.1. Methods

A summary [15] of our focal data set provides all variable definitions, sampling strategies, and code for analyzing experimental data in a variety of ways. Data sets and images of experimental conditions are also available online [14] and at `https://github.com/bgpurzycki/Moral-Models-Moral-Behavior`. The following variables represent a distillation of models published elsewhere [16, 18]:

1. **Moralistic gods' punishment** was the composite mean value of two questions with binary response values (no = 0; yes = 1): *Does* [DEITY NAME] *ever punish people for their behavior?* and *Can* [DEITY NAME] *influence what happens to people after they die?*
2. **Moralistic gods' knowledge** was the mean value of two binary-response questions (no = 0; yes = 1) about the breadth of deities' knowledge: *Can* [DEITY NAME] *see into peoples hearts or know their thoughts and feelings?* and *Can* [DEITY NAME] *see what people are doing if they are far away in [a distant town or city familiar to locals]?*
3. **Number of children** was self-reported number of children "fathered or given birth to."
4. **Treatment condition** is an indicator variable denoting whether (= 1) or not (= 0) participants played games in the presence of an image or object with local religious significance. This was a manipulation in which these objects were selected on the basis of having no visual indices of agency (i.e., no eyes).
5. **Game order** is an indicator variable tracking which game participants played first. This is denoted with a "1" if participants played the Local Community Game first and a "0" if they played the Self Game first.
6. **Game was about honesty?** denotes whether responses to post-experimental questions asking participants what they thought the games were about included "honesty," "fairness," and/or "cheating."

Recall from the main text that some sites had religious prime conditions. These sites were: Lovu (Fiji), Mauritius, Marajó, Tyva Republic, and Yasawa. Lovu Fijians in the treatment condition played games on a table with a statue that abstractly represented Shiva (a *lingam*). Mauritians in the treatment condition played in a Hindu temple. Brazilian participants played near a Bible and crucifix necklace, the Tyvan condition included playing near a Buddhist luck charm (*kamgalal*), and Yasawans in the treatment condition played on a cloth with a cross, Bible, and Bible verse (Mark 9:23) printed on it. Further descriptions and images of prime conditions can be found at: `https://github.com/bgpurzycki/Evolution-of-Religion-and-Morality`. Again, we found no overall relationship between playing in treatment conditions and game outcome. We nevertheless hold this factor constant in our models.

Note that the Hadza were not asked what they thought the game was about. In the main models, we imputed this and all other missing data from prior distributions (see Supplementary section 3.2.3. for distributions and R script for implementation). For the frequentist models, we used a couple of different imputation techniques (see below).

4.2. Main model specifications

The statistical analysis is essentially a binomial regression with varying effects used to manage repeat observations on individuals and groups. The additional feature of the model is that we use group-level predictors to express average norms and beliefs in each group. However, these predictors are not directly observable and must be estimated themselves. Therefore, simultaneous models are run to: (1) estimate those predictors, and (2) plug the resulting variable distributions directly into the individual-level regression. This method retains all uncertainty arising from sample size and variance in each group. Below, we explain the full model in pieces.

4.2.1. Coin model

To model the coin allocations y out of 30, we use a binomial regression:

$$y_i \sim \text{Binomial}(30, p_i) \tag{1}$$

$$\text{logit}(p_i) = \alpha + \sigma_{\text{id}} z_{\text{id[i]}} + a_{\text{group[i]}} \tag{2}$$

$$+ b^h_{\text{group[i]}} h_i + b^H H_{\text{group[i]}} \tag{3}$$

$$+ b^p_{\text{group[i]}} p_i + b^P P_{\text{group[i]}} \tag{4}$$

$$+ b^o_{\text{group[i]}} o_i + b^O O_{\text{group[i]}} \tag{5}$$

$$+ b^{\text{children}} k_i + b^{\text{condition}} t_i + b^{\text{order}} r_i + b^{\text{check}} \chi_i + b^{\text{game}} g_i \tag{6}$$

The linear predictor $\text{logit}(p_i)$ measures partial associations between the amount allocated to the distant cup and both group and individual variables. Line (2) includes an intercept α for the full sample and varying intercepts for individual and group[1]. The next three lines express the three cultural variables of interest, at both individual and

[1]Efficient sampling is made possible by the use of the non-centered prior for the individual varying effects, the $\sigma_{\text{id}} z_{\text{id[i]}}$ construction in the coin model where z_{id} represents the varying effect by individual; $z_{id} \sim \text{Normal}(0, 1)$, $\sigma_{id} \sim \text{Exponential}(1)$. We attempted the model first with a traditional centered parameterization, where the scale parameter σ_{id} was in the prior for z_{id}, but that made analyses fail. We therefore put it directly in the linear model.

group levels: honesty (h/H), gods' punishment (p/P) and knowledge breadth (i.e., omniscience) (o/O). Individual responses are given lowercase letters whereas group averages are given capital letters. Each group has its own varying slope for individual variables, allowing the relationship between individual response (e.g., whether they listed (dis)honesty) and behavior to vary across groups. The last line, (6), contains simple (i.e., fixed) effects for number of children, the experimental treatment condition, game order, game check, and an indicator for game (Self Game = 1), respectively.

All simple effects above are assigned weakly-regularizing Normal$(0, 1)$ priors. These guard against finding strong effects in small samples or those that vary considerably in responses, but are easily overwhelmed in large or consistent samples.

The varying intercepts for individuals are given a prior scale of:

$$\sigma_{\text{id}} \sim \text{Exponential}(1)$$

This is likewise weakly regularizing. The varying effects for group are bound together in a common variance-covariance matrix:

$$\begin{bmatrix} a_j \\ b_j^h \\ b_j^p \\ b_j^o \end{bmatrix} \sim \text{MVNormal}\left(\begin{bmatrix} 0 \\ \beta^h \\ \beta^p \\ \beta^o \end{bmatrix}, \mathbf{SRS} \right)$$

where $\mathbf{S}$ is a diagonal matrix of standard deviations of the intercept and the three cultural variables of interest:

$$\mathbf{S} = \begin{bmatrix} \sigma_a & 0 & 0 & 0 \\ 0 & \sigma_{b^h} & 0 & 0 \\ 0 & 0 & \sigma_{b^p} & 0 \\ 0 & 0 & 0 & \sigma_{b^o} \end{bmatrix}$$

and $\mathbf{R}$ is a full rank correlation matrix of the same variables. Each standard deviation is assigned an independent Exponential(1) prior as before, and $\mathbf{R}$ is given a weakly regularizing prior from the LKJ family [10] as implemented by RStan 2.14.1 [20]:

$$\mathbf{R} \sim \text{LKJCorr}(4)$$

This is necessary because, unlike the four standard deviations, the six individual correlations inside $\mathbf{R}$ cannot be independent of one another. The constraint that the correlation matrices be positive definite severely constrains possible combinations of the six parameters. The LKJCorr family manages this by composing symmetric Beta distribution shapes, that is, symmetric distributions of potential correlations. As the dimension of such a matrix and the shape parameter η grow, the prior is increasingly concentrated on the identity matrix (i.e., all numbers on the diagonal are 1 and all other values are 0). The value we use here, $\eta = 4$, regularizes by penalizing extreme correlations (e.g., 1 or -1). Sometimes strong regularization is needed to fit such a model, but we found the model sampled just fine even with no regularization imposed on this matrix. We retain the regularization because it is best practice.

4.2.2. Group belief models

Three of the predictor variables above are unobserved: H_{group}, P_{group}, and O_{group}. These are average honesty, punishment, and omniscience beliefs for each group. Since these cannot be observed, we infer them from the sample of individual statements in that group. We simultaneously fit these three models and the coin model above using the posterior distribution of each in the coin model. This retains all uncertainty so that we do not impart false precision to the estimates.

In each of the three cases, we treat the unobserved variables as varying intercepts for each group. For example, the honesty model is:

$$\begin{aligned} h_i &\sim \text{Binomial}(2, p_i) \\ \text{logit}(p_i) &= H_{\text{group[i]}} \\ H_j &\sim \text{Normal}(\mu^H, \sigma^H) \\ \mu^H &\sim \text{Normal}(0, 5) \\ \sigma^H &\sim \text{Exponential}(1) \end{aligned}$$

The observed h_i values take the values 0, 1, or 2, where the values represent neither, one, or both honesty and dishonesty were listed. The above is therefore a binomial GLMM with two trials. The latent estimate H_{group} is our target of inference, and H_j defines the priors for each group. The models for P_{group} and O_{group} are constructed analogously.

4.2.3. Imputation models

For the main model, we relied on the following distributions for the imputation of missing data:

$$\begin{aligned}
\text{children} &\sim \text{Normal}(\mu_{\text{children}}, \sigma_{\text{children}}) \\
\mu_{\text{children}} &\sim \text{Normal}(1, 1) \\
\sigma_{\text{children}} &\sim \text{Exponential}(10) \\
\text{order} &\sim \text{Bernoulli}(0.5) \\
\text{check} &\sim \text{Bernoulli}(\phi_{\text{check}}) \\
\phi_{\text{check}} &\sim \text{Beta}(1, 1) \\
\sigma_{\text{punishment}} &\sim \text{Exponential}(1) \\
\sigma_{\text{omniscience}} &\sim \text{Exponential}(1)
\end{aligned}$$

4.2.4. Sampling and diagnostics

We took 500 samples from each of the 4 chains with a target acceptance rate of 0.99 and all other settings the default as of RStan 2.14.1. Sampling was very efficient. All $\hat{R}$ values were below 1.01 and effective sample sizes were greater than 50% of the actual sample count.

5. Supplementary Analyses

In addition to the primary Bayesian models in the main text and above, we also ran a variety of model specifications using standard multi-level binomial logistic regressions for the sake of comparison and robustness checking. All models were fit using the lme4 package [3] for use in R.

Note again that the Hadza were not asked what they thought the game was about (the "game check" variable discussed above). In these models, we imputed these missing values in two ways. One method relied on an imputation function that randomly selects data points from the rest of the sample [7, p. 534]. These results are reported here. We also used a dummy value for the missing Hadza game check data and also ran the same models without the Hadza, and the qualitative results hold (results are not reported here, but procedures are included in the R script). In general the individual-level results are qualitatively the same as the main model, but we do highlight some important modeling differences throughout the discussion.

Table S4 reports full models, all of which vary gods' punishment and knowledge breadth across sites. They also vary moral model values (i.e., summation of (dis)honesty, and the salience of honesty and dishonesty). Table S5 reports results from models that tease apart the specific components of the free-list data and assesses the role that, for instance, listing honesty or dishonesty individually play. As participants played two games, all models have varying intercepts for individuals.

5.1. Varying cultural variables across sites

Table S4 reports five model specifications. The first is a reduced model that only includes moralistic deities' knowledge and punishment as varied effects across sites. The second model (model mS4.full in the R script) includes varying effects for the three cultural variables—summation of (dis)honesty, moralistic gods' punishment and knowledge breadth—across sites; each site has its own intercept and each effect has its own slope. The next model (model mS4.1) varies all of the cultural variables independently by site, where site has only a single intercept associated with (dis)honesty. The next two models are the same, but instead of the summations of (dis)honesty, we consider cognitive salience.

Does cognitive salience or accessibility of (dis)honesty predict honest play? We would expect that the degree to which individuals can access task-relevant components to moral models will also play a role in motivating individuals to behave in accordance with their moral prescriptions. In order to test this, we assume that for all instances where participants' free-lists lack honesty and/or dishonesty, salience is zero. Because *not* listing these items translates to

Predictor	**No Moral Model** OR [95% CI]	**Sum of (Dis)honesty** OR [95% CI]	**Sum of (Dis)honesty** OR [95% CI]	**Honesty (Sal.)** OR [95% CI]	**Dishonesty (Sal.)** OR [95% CI]
Moral model variable	—	1.05 [0.98, 1.13]	1.05 [0.99, 1.12]	1.11 [0.93, 1.33]	1.12 [0.99, 1.28]
Moralistic gods' punishment	1.15 [1.02, 1.29]	1.15 [1.02, 1.29]	1.14 [1.03, 1.28]	1.15 [1.03, 1.28]	1.14 [1.02, 1.27]
Moralistic gods' knowledge	1.30 [1.02, 1.65]	1.34 [1.04, 1.72]	1.23 [1.04, 1.45]	1.24 [1.05, 1.46]	1.22 [1.04, 1.42]
Number of children	1.00 [0.98, 1.01]	1.00 [0.98, 1.01]	1.00 [0.98, 1.01]	1.00 [0.98, 1.01]	1.00 [0.98, 1.01]
Condition (Prime = 1)	0.98 [0.91, 1.06]	0.98 [0.91, 1.06]	0.97 [0.90, 1.05]	0.97 [0.90, 1.05]	0.97 [0.90, 1.05]
Local Game played first = 1	1.04 [0.97, 1.11]	1.04 [0.97, 1.11]	1.04 [0.98, 1.11]	1.04 [0.98, 1.11]	1.05 [0.98, 1.12]
Game about honesty? (Yes = 1)	0.94 [0.82, 1.08]	0.93 [0.81, 1.07]	0.93 [0.81, 1.07]	0.93 [0.81, 1.07]	0.93 [0.81, 1.07]
Game (Self Game = 1)	0.98 [0.94, 1.03]	0.98 [0.94, 1.03]	0.98 [0.94, 1.03]	0.98 [0.94, 1.03]	0.98 [0.94, 1.03]
Intercept	0.63 [0.47, 0.83]	0.60 [0.44, 0.81]	0.65 [0.54, 0.78]	0.65 [0.54, 0.78]	0.66 [0.55, 0.79]
log likelihood	-2727.3	-2725.2	-2727.0	-2727.4	-2726.2
AIC	5484.7	5490.4	5484.1	5484.8	5482.3
Model convergence?	no	no	no	no	no
Model name in R script	mS4.red	mS4.full	mS4.1	mS4.2	mS4.3

Table S4: **Exponentiated coefficients (OR) and 95% confidence intervals (CI) of simple effects for full models varying gods' punishment, knowledge breadth, and moral model variables by site.** Target moral model simple and varying effects are listed across the top row where "Sal." refers to individual salience of honesty or dishonesty. Across all models, $N = 508$.

missing values in data processing, we replaced these missing values with zero for all participants who did free-lists tasks, but did not list honesty or dishonesty.

Again, as these lists were intended to be capped at a maximum of 5 items listed, these results should be interpreted with caution. Moreover, recall that in some sites, few individuals listed (dis)honesty (see Table 5 in main text). This may create by-site interpretations difficult. The R script nevertheless provides code for examining these differences. Using the same model as mS4.1, but instead using item salience to predict game outcome, the order in which individuals list honesty or dishonesty has an effect; on average, listing these items earlier predicts greater odds of allocating coins to the distant cup. Note, however, that the effect is both better estimated and stronger for the salience of dishonesty (mS4.3; and the model properly converged) than honesty (mS4.2). If this effect is real, it suggests that how salient or accessible these items are matters in game outcome as well. In other words, *how quick* individuals are to equate good and bad people with such virtues when prompted also plays an important role on human behavior. With the aforementioned caveats in mind, these results suggest that conceptual salience matters in the production of moral behavior.

These models failed to converge and have removed all cases with missing data. Nevertheless, the results suggest that moral models, gods' punishment and knowledge breadth increase the odds that players put a coin into the cup reserved for geographically distant co-ethnic, co-religionists, qualitatively similar to the results reported in the main text.

5.2. Varying only moral models across sites

5.2.1. Model specifications

Table S5 reports a variety of model specifications designed to examine in the role honesty and dishonesty specifically play as individual variables. Model 1 is based on a reduced model reported elsewhere [18]. Its main differences from those models are: (1) here we use the imputed game check values for the Hadza, and (2) we ignore group-level variation (i.e., it is a fully pooled model ignoring groups, but does vary intercepts for individuals). Model 2 is the same basic model, but varies the intercept for individuals *and* field sites. Note that while the qualitative effects are the same, there is a slight decrease in the odds ratio and a shift in the range of the confidence intervals for the influence of moralistic gods' punishment. This further emphasizes the importance of the effects of cross-site variation in gods' punishment values have on game outcome.

Model 3 builds on Model 2 by adding the summation of listing "honesty" and "dishonesty." The model varies the intercepts for both individuals (as the two game outcomes are repeated measures) and field sites. Though slight, the content of individuals' moral models does have an effect on allocations in the predicted direction; when individuals report that "honesty" and "dishonesty" are indicative of moral or immoral people respectively, they have a greater chance of allocating a coin to the distant co-ethnic, co-religionist. Though relatively slight, the effect remains stable across model specifications.

Model 4 builds on Model 3 by varying the slopes by site for the honesty and dishonesty summation values. We reasoned that when more people in these sites list honesty and dishonesty, individual allocations should more likely go to the geographically distant players. In other words, individuals from communities where honesty and dishonesty are more salient moral values should play more fairly. As such, the effect of group-level honesty and dishonesty should vary across sites. The estimates of all simple effects are consistent with the other models.

Models 5-7 tease apart the free-list summations and include only "honesty" (Model 5), "dishonesty" (Model 6), and both as individual variables (Model 7). Across models, when someone lists honesty and/or dishonesty, the chances of allocating a coin to the distant cup increases by $\sim$5%, the highest upper bound of the confidence interval across these models was 19%.

5.2.2. Group-level projections

Based on Model 4 in Table S5, Figure S1 models the cross-site effects of listing honesty and dishonesty on the probability of allocating a coin to the distant player. Baseline trends (modeled by the gray lines) are the logistic transformed summations of the simple intercept coefficient (-0.42) and the coefficient for moral models (0.05) times the possible moral model values (0, 1, or 2, therefore 0.40, 0.41, and 0.42 respectively). Hollow points are logistic transformed summations of the simple intercept coefficient (-0.43), the site-specific intercept, the simple, individual-level effect on moral models (0.05), and the by-site slope coefficients for moral models times their possible values of 0, 1, or 2. Defined in this fashion, these points model the group-level contributions of moral models: when x-axis values are zero, this is akin to modeling the individual-level probability of giving to a distant individual when group members mentioned *neither* honesty nor dishonesty on average; when x-axis values are two, this is akin to modeling the probability of giving a coin to the distant player if group-level moral models were to include *both* honesty and

Predictor	**Model 1** OR [95% CI]	**Model 2** OR [95% CI]	**Model 3** OR [95% CI]	**Model 4** OR [95% CI]	**Model 5** OR [95% CI]	**Model 6** OR [95% CI]	**Model 7** OR [95% CI]
Honesty (Present = 1)	—	—	—	—	1.05 [0.94, 1.19]	—	1.05 [0.94, 1.17]
Dishonesty (Present = 1)	—	—	—	—	—	1.06 [0.97, 1.16]	1.05 [0.96, 1.15]
(Dis)honesty summation	—	—	1.04 [0.98, 1.09]	1.05 [0.98, 1.12]	—	—	—
Moralistic gods' punishment	1.23 [1.10, 1.37]	1.15 [1.03, 1.28]	1.15 [1.03, 1.28]	1.15 [1.03, 1.28]	1.15 [1.03, 1.28]	1.14 [1.02, 1.27]	1.15 [1.03, 1.28]
Moralistic gods' knowledge	1.20 [1.02, 1.40]	1.22 [1.04, 1.42]	1.21 [1.04, 1.42]	1.22 [1.04, 1.43]	1.22 [1.04, 1.43]	1.22 [1.04, 1.43]	1.22 [1.04, 1.43]
Number of children	0.99 [0.98, 1.01]	1.00 [0.98, 1.01]	1.00 [0.98, 1.01]	1.00 [0.98, 1.01]	1.00 [0.98, 1.01]	1.00 [0.98, 1.01]	1.00 [0.98, 1.01]
Condition (Prime = 1)	0.99 [0.92, 1.07]	0.98 [0.90, 1.05]	0.97 [0.90, 1.05]	0.97 [0.90, 1.05]	0.97 [0.90, 1.05]	0.97 [0.90, 1.05]	0.97 [0.90, 1.05]
Local Game played first = 1	1.05 [0.98, 1.12]	1.04 [0.98, 1.12]	1.04 [0.98, 1.12]	1.04 [0.98, 1.12]	1.04 [0.98, 1.11]	1.05 [0.98, 1.12]	1.04 [0.98, 1.12]
Game about honesty? (Yes = 1)	0.90 [0.78, 1.04]	0.94 [0.82, 1.08]	0.94 [0.81, 1.08]	0.93 [0.81, 1.07]	0.93 [0.81, 1.08]	0.94 [0.81, 1.08]	0.93 [0.81, 1.07]
Game (Self Game = 1)	0.98 [0.94, 1.03]	0.98 [0.94, 1.03]	0.98 [0.94, 1.03]	0.98 [0.94, 1.03]	0.98 [0.94, 1.03]	0.98 [0.94, 1.03]	0.98 [0.94, 1.03]
Intercept	0.66 [0.56, 0.78]	0.67 [0.56, 0.80]	0.66 [0.55, 0.79]	0.65 [0.54, 0.79]	0.66 [0.55, 0.79]	0.66 [0.55, 0.79]	0.65 [0.54, 0.79]
Vary intercept for field site?	no	yes	yes	yes	yes	yes	yes
Vary slope for free-list value?	no	no	no	yes	yes	yes	yes*
log likelihood	-2738.6	-2728.6	-2727.8	-2727.1	-2728.0	-2727.4	-2727.1
AIC	5495.3	5477.1	5477.6	5480.2	5482.0	5480.8	5482.2
Model convergence?	yes	yes	yes	no	yes	no	no
Model name in R script	mS5.1	mS5.2	mS5.3	mS5.4	mS5.5	mS5.6	mS5.7

Table S5: **Odds ratios and 95% confidence intervals for hierarchical binomial logistic regressions for both games.** In all models, intercepts varied by participant ($N = 494$). Varied slopes for free-list values are the same as corresponding simple effect except for *, where the summation value was used (as in Models 3 and 4). We imputed Hadza values for the game check ("Game about honesty?") for all models.

dishonesty. Solid points indicate where the site-specific mean of (dis)honesty summations are along the x-axis, placed on the regression line for reference.

Two sites allocate coins far below sample-baseline allocations, the Hadza and Yasawans. Lovu Fijians and Tyvans largely play fairly, as indicated by their values closer to 0.5 and having greater-than-baseline allocations overall. As illustrated in Figure S1, the probability of Yasawan and Hadza allocations to the distant player would dramatically increase as a function of increasing the group-level ubiquity of honesty and dishonesty in moral models. The Lovu Fijians and Tyvans show barely noticeable effects, but they remain positive. In other words, compared to contexts with low probabilities of allocating coins to distant players and infrequent cases of listing (dis)honesty, extrapolating from this model adds little in the way of projection for contexts where honesty and dishonesty are already prevelent. Nevertheless, it does show that increasing prevelence of (dis)honesty can have an impact on game play.

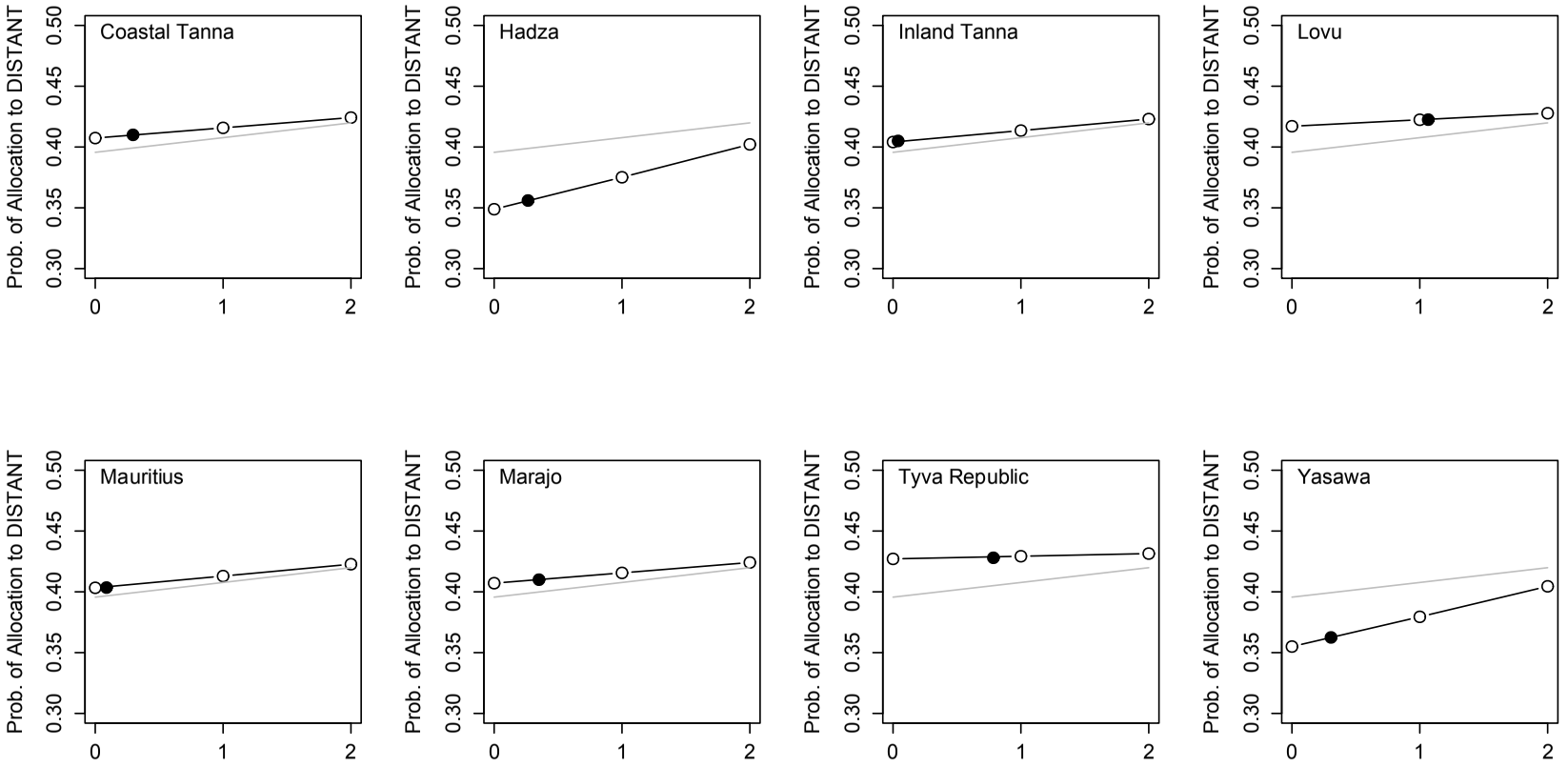

Figure S1: **Projected group-level effects of moral models on the probability of allocating a coin to geographically distant players across eight field sites.**. Projections are generated from Model 4 in Table S5. X-axes are the possible summation values of the presence of honesty and dishonesty in free-list tasks (0, 1, or 2). Values of 0.50 on y-axes indicate an unbiased allocation. Points are logistic transformed summations of the simple intercept coefficient, the site-specific intercept, the simple effect on moral models, and the by-site varied slope coefficients for moral models times their possible values of 0, 1, or 2. Gray lines are the baseline effects of moral models on the probability of allocating a coin to distant players (the same across all plots). Black lines indicate direction and magnitude of projected effect of increasing ubiquity of "honesty" and "dishonesty" on game play in each population. Solid points indicate location of site-specific mean of the (dis)honesty summations placed on the regression line.

6. References

[1] Apicella, C. L., 2018. High levels of rule-bending in a minimally religious and largely egalitarian forager population. Religion, Brain & Behavior 8 (2), 133–147.

[2] Atkinson, Q. D., 2018. Religion and expanding the cooperative sphere in Kastom and Christian villages on Tanna, Vanuatu. Religion, Brain & Behavior 8 (2), 149–167.

[3] Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67 (1), 1–48.

[4] Bonnemaison, J., 1994. The Tree and the Canoe: History and Ethnogeography of Tanna. University of Hawaii Press, Honolulu.

[5] Carmeli, Y. S., Eriksen, T. H., 1998. Common Denominators: Ethnicity, Nation-Building and Compromise in Mauritius, first edition edition Edition. Bloomsbury Academic, New York.

[6] Cohen, E., Baimel, A., Purzycki, B. G., 2018. Religiosity and resource allocation in Marajó, Brazil. Religion, Brain & Behavior 8 (2), 168–184.

[7] Gelman, A., Hill, J., 2007. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, Cambridge.

[8] Gervais, M. M., 2013. Structures of Sentiment: Mapping the Affective Bases of Social Relationships in Yasawa, Fiji. University of California, Los Angeles.

[9] Lal, B. V., 1992. Broken Waves: A History of the Fiji Islands in the Twentieth Century. University of Hawaii Press, Honolulu.

[10] Lewandowski, D., Kurowicka, D., Joe, H., 2009. Generating random correlation matrices based on vines and extended onion method. Journal of Multivariate Analysis 100, 1989–2001.

[11] Marlowe, F., 2010. The Hadza: Hunter-gatherers of Tanzania. Univ of California Press.

[12] McNamara, R. A., Norenzayan, A., Henrich, J., 2016. Supernatural punishment, in-group biases, and material insecurity: experiments and ethnography from Yasawa, Fiji. Religion, Brain & Behavior 6 (1), 34–55.

[13] Purzycki, B. G., Apr. 2016. The evolution of gods' minds in the Tyva Republic. Current Anthropology 57 (S13), S88–S104.

[14] Purzycki, B. G., Apicella, C., Atkinson, Q. D., Cohen, E., Henrich, J., McNamara, R. A., Norenzayan, A., Willard, A. K., Xygalatas, D., 2016. Evolution of Religion and Morality Project Dataset (Wave 1). Harvard Dataverse.
URL `https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/RT5JTV`

[15] Purzycki, B. G., Apicella, C., Atkinson, Q. D., Cohen, E., McNamara, R. A., Willard, A. K., Xygalatas, D., Norenzayan, A., Henrich, J., 2016. Cross-cultural dataset for the evolution of religion and morality project. Scientific Data 3, 160099.

[16] Purzycki, B. G., Apicella, C., Atkinson, Q. D., Cohen, E., McNamara, R. A., Willard, A. K., Xygalatas, D., Norenzayan, A., Henrich, J., 2016. Moralistic gods, supernatural punishment and the expansion of human sociality. Nature 530 (7590), 327–330.

[17] Purzycki, B. G., Kulundary, V., 2018. Buddhism, identity, and class: fairness and favoritism in the Tyva Republic. Religion, Brain & Behavior 8 (2), 205–226.

[18] Purzycki, B. G., Norenzayan, A., Apicella, C., Atkinson, Q. D., Cohen, E., Henrich, J., McNamara, R. A., Willard, A. K., Xygalatas, D., 2018. The evolution of religion and morality: A synthesis of ethnographic and experimental evidence from eight societies. Religion, Brain & Behavior 8 (2), 100–132.

[19] Smith, K. D., Smith, S. T., Christopher, J. C., 2007. What defines the good person? cross-cultural comparisons of experts' models with lay prototypes. Journal of Cross-Cultural Psychology 38 (3), 333–360.

[20] Stan Development Team, 2016. RStan: the R interface to Stan. R package version 2.14.1.
URL http://mc-stan.org/

[21] Vainshtein, S., et al., 1980. Nomads of South Siberia: The Pastoral Economies of Tuva. Cambridge University Press, Cambridge.

[22] Willard, A. K., 2018. Religion and prosocial behavior among the Indo-Fijians. Religion, Brain & Behavior 80 (2), 227–242.

[23] Xygalatas, D., Kotherová, S., Maňo, P., Kundt, R., Cigán, J., Klocová, E. K., Lang, M., 2018. Big gods in small places: The random allocation game in Mauritius. Religion, Brain & Behavior 8 (2), 243–261.