Journal of Economic Behavior and Organization 215 (2023) 134–152

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

The behavioral mechanisms of voluntary cooperation across culturally diverse societies: Evidence from the US, the UK, Morocco, and Turkey

Till O. Weber [a,*], Jonathan F. Schulz [b], Benjamin Beranek [c], Fatima Lambarraa-Lehnhardt [d], Simon Gächter [e,f,g,*]

[a] Newcastle University Business School, Newcastle University, UK
[b] Department of Economics, George Mason University, USA
[c] Department of Finance, Economics and Risk Management, USA
[d] Leibniz Centre for Agricultural Landscape Research (ZALF), Müncheberg, Germany
[e] School of Economics, University of Nottingham, UK
[f] IZA Bonn, Germany
[g] CESifo Munich, Germany

ARTICLE INFO

JEL:
C9
H4
C7
D2

Keywords:
Public goods
Voluntary cooperation
ABC method
Conditional cooperation
Beliefs
Punishment
Cross-cultural experiments
WEIRD societies

ABSTRACT

We examine the role of cooperative preferences, beliefs, and punishments to uncover potential cross-societal differences in voluntary cooperation. Using one-shot public goods experiments in four comparable subject pools from the US and the UK (two similar Western societies) and Morocco and Turkey (two comparable non-Western societies), we find that cooperation is lower in Morocco and Turkey than in the UK and the US. Using the ABC approach – in which cooperative attitudes and beliefs explain cooperation – we show that cooperation is mostly driven by differences in beliefs rather than cooperative preferences or peer punishment, both of which are similar across the four subject pools. Our methodology is generalizable across subject pools and highlights the central role of beliefs in explaining differences in voluntary cooperation within and across culturally, economically, and institutionally diverse societies. Because our behavioral mechanisms correctly predict actual contributions, we argue that our approach provides a suitable methodology for analyzing the determinants of voluntary cooperation of any group of interest.

1. Introduction

Trust and social capital are crucial for economic prosperity, but vary considerably across societies (e.g., Knack and Keefer (1997); Guiso et al. (2008); Tabellini (2008); Fehr (2009); Algan and Cahuc (2013)). Social capital is often expressed as voluntary cooperation in situations where collective welfare and self-interest are in conflict. Here, we take a cross-societal perspective and use laboratory public goods experiments in four countries to investigate two behavioral mechanisms that likely generate variation in voluntary cooperation: (i) conditional cooperation and beliefs, and (ii) punishment.

An important finding from years of behavioral research on voluntary cooperation is that many people are *conditional cooperators*:

* Corresponding authors.
E-mail addresses: tilloweber@gmail.com (T.O. Weber), simon.gaechter@nottingham.ac.uk (S. Gächter).

https://doi.org/10.1016/j.jebo.2023.09.006
Received 8 April 2022; Received in revised form 29 August 2023; Accepted 2 September 2023
Available online 21 September 2023

their own willingness to cooperate increases with the cooperation of others.[1] A second important (and related) finding is that *beliefs* about others' contributions are positively correlated with people's own contributions (e.g., Croson (2007); Fischbacher and Gächter (2010); Gächter and Renner (2018)). To control for the problem of belief endogeneity, Fischbacher et al. (2001) introduced an incentive-compatible method to elicit people's cooperative attitudes, that is, their willingness to cooperate as a function of all levels of cooperation of others. Using this method, Fischbacher and Gächter (2010), Fischbacher et al. (2012), Gächter et al. (2017), Isler et al. (2021), and Gächter et al. (2022) found that cooperative *a*ttitudes (a) and *b*eliefs (b) together explain the observed level of voluntary *c*ooperation (c), i.e., $a(b)\rightarrow c$. Gächter et al. (2017) refer to this approach as the 'ABC of cooperation', which is our first behavioral mechanism. The ABC methodology enables us to see whether the societal background affects a or b, or both, such that it explains the observed societal-level variation in cooperation c. If differences in societal background matter for how people play a one-shot public goods game with identical incentives in each subject pool, then the ABC method should be able to pick it up and pinpoint where differences in cooperation, should they exist, are coming from.

A further robust result is that many people are willing to incur costs to *punish* group members even in one-shot public goods games where a first-stage cooperation game is followed by possibilities of costly peer punishment at the second-stage (e.g., Fehr and Gächter (2002); Walker and Halloran (2004); Casari (2005); Anderson and Putterman (2006); Carpenter (2007); Egas and Riedl (2008); Nikiforakis and Normann (2008); Gächter and Herrmann (2011); Weber et al. (2018); Molleman et al. (2019)).[2] Punishment of free riding thus provides a further channel affecting cooperative choices that is independent of cooperative attitudes and cooperative beliefs (Weber et al. (2018)). This can be the case if punishment is a credible threat. Punishment is a credible threat if people believe that too low a contribution might trigger sufficiently strong punishment from other group members (for evidence on this see, e.g., the survey by Gächter and Herrmann (2009)). In this case cooperating becomes a dominant strategy that is independent of one's cooperative attitude and belief in the underlying (first stage) public goods game. Therefore, punishment provides a further reason to cooperate – punishment is our second behavioral mechanism.

With a few exceptions, which we review in Section 2, most results on the two behavioral mechanisms we study here come from the US and Western European subject pools, that is, from societies that Henrich et al. (2010) have labelled WEIRD: Western, Educated, Industrialized, Rich and Democratic (see also Henrich (2020) and Apicella et al. (2020)). How generalizable are these findings to culturally more diverse subject pools? In this paper, we approach this question by using the ABC methodology to compare conditionally cooperative attitudes, beliefs, and punishment in two Western societies—the classically WEIRD societies of the US and the UK—and two non-Western societies—Morocco and Turkey—that are culturally among the most distant countries from the US and the UK in a quantifiable way (based on Muthukrishna et al. (2020), the most comprehensive dataset on cultural distance available; see Section 3.2). We therefore provide a methodological contribution in testing the generalizability of the ABC framework across culturally more diverse subject pools than used in previous research.

Our goal is to go beyond simply observing whether there are differences in voluntary cooperation levels in different societies (there are, see, e.g., Herrmann et al. (2008); Gächter and Herrmann (2009)). Instead, with the ABC methodology we quantitatively measure several factors that plausibly underpin voluntary cooperation and assess the predictive power of our two behavioral mechanisms to explain cooperation. Choosing subject pools with comparable age, education, and socio-economic status in their respective societies (students) allows us to attribute sample differences in our variables of interest to cultural variation while minimizing confounds. Our approach opens the "black box" of voluntary cooperation because it allows us to pinpoint where in the behavioral mechanisms (attitudes, beliefs, and/or punishment) cross-societal differences in voluntary cooperation occur—if at all, because the incentives are of course identical in all subject pools.[3]

Our paper thus does not aim to investigate how historical, cultural, or institutional factors causally impact voluntary cooperation (for an example of this line of research, see Rustagi (2022)). Our goal is instead to contribute to the literature that aims to better understand the mechanisms of how cultural differences might influence economic outcomes (e.g., Guiso et al. (2006); Alesina and Giuliano (2015))—in our case levels of voluntary cooperation. This literature emphasizes the role of both beliefs and preferences as distinct aspects of culture (see e.g., Greif (1994) on the role of cultural beliefs). Yet, most empirical papers do not, or are unable to, distinguish between beliefs and preferences (see Alesina and Giuliano (2015) for a discussion). By using the ABC approach, we shed light on the relative importance of cooperative attitudes (as a proxy for preferences) and beliefs in accounting for potential differences in voluntary cooperation in comparable subject pools from countries that vary substantially in their societal background.

We use anonymous laboratory one-shot public goods games to study the behavioral mechanisms of ABC and punishment. The one-shot nature of the public goods game allows us to measure the elements of the two behavioral mechanisms without having to deal with potential confounds arising from repeated play. Both mechanisms consist of a sequence of incentivized experiments that all partici-

[1] In lab experiments, e.g., Keser and van Winden (2000); Fischbacher et al. (2001); Gächter et al. (2017); Gächter et al. (2022); Bilancini et al. (2022); in lab-in-the-field experiments, e.g., Rustagi et al. (2010); in experiments with representative samples, e.g., Thöni et al. (2012); Fosgaard et al. (2014); in field experiments, e.g., Frey and Meier (2004). See also Gächter (2007) and Thöni and Volk (2018) for reviews.

[2] For reviews see Gächter and Herrmann (2009); Chaudhuri (2011); Fehr and Schurtenberger (2018); Raihani and Bshary (2019).

[3] In using this approach, we anticipate fewer differences in the relative education and income across our four student samples than one would expect when comparing samples drawn from the respective general populations. Fully representative samples also require sample sizes that were infeasible for us to run in a controlled laboratory setting. Because we could only run small samples in the lab, and an important element of our contribution is methodological, we opted for a controlled demographic. See Gächter (2010) for a discussion on using student subject pools. Of course, if large and representative samples are feasible, then demographic variation might result in additional interesting variation across our key ABC variables.

pants play in the same order. To test the first behavioral mechanism (the ABC approach), participants play a one-shot public goods game with no punishment (called the "N-Game") in which we elicit two types of decisions for each participant *i*: (*i*) an unconditional contribution c_i to the public good, and (*ii*) contribution decisions that are a function $a_i(\overline{c}_j),\ i \neq j$, of all possible rounded average contribution $\overline{c}_j$ of the *j* other group members. We also elicit *beliefs* b_i (as a point estimate) about the average contribution $\overline{c}_j$ of the *j* other group members. Based on previous findings (Fischbacher and Gächter (2010); Gächter et al. (2017); Isler et al. (2021); and Gächter et al. (2022)), we predict that cooperative attitudes and beliefs—both of which might be subject-pool specific—jointly determine cooperation: $a_i(b_i) = \widehat{c}_i \approx c_i$, that is, the contribution predicted by the ABC approach, $\widehat{c}_i$, corresponds to the observed contribution c_i, regardless of potential differences in a_i and b_i between subject pools.

For the second behavioral mechanism (punishment), participants play a first-stage one-shot public goods game followed by a second stage with peer punishment (called the "P-game"). We focus on preferences for *enforcement of cooperation* by studying how each participant's expected punishment affects their contribution decision. The one-shot design removes strategic incentives to engage in punishment as well as the possibility of feuds and spill-over effects over rounds (e.g., Nikiforakis (2008); Nikiforakis and Engelmann (2011)). Any willingness to cooperate in the N-Game or pay for punishment in the P-Game cannot be motivated by strategic concerns and is therefore reflecting a "preference".

Our main findings are as follows. Overall, most people have conditionally cooperative attitudes in all subject pools, and the differences between them are small, i.e., 'WEIRDness' has little effect on cooperative attitudes. The same is true for antisocial punishment. The main difference between the subject pools is in beliefs: In both experiments, British and American subjects are more optimistic about the contributions of others and therefore contribute more to the common good than Moroccan and Turkish subjects. When punishment is available, UK and US participants are more likely to expect punishment for free riding.

We have two main conclusions. First, the subject pools we compare here were chosen to span culturally, economically, and institutionally distant countries. Despite these substantial *ex ante* differences of the societies of our subject pools, our empirical findings paint a consistent picture of the underlying behavioral mechanisms: attitudes to conditional cooperation (*a*) show only small variation, beliefs (*b*) are different between subject pools, and differences in overall cooperation levels (*c*) across subject pools are mostly explained by differences in beliefs. The ABC approach can therefore explain cooperation levels in all subject pools. Because in the experiment participants do not know who they interact with, and because the experiment is one-shot and anonymous, the beliefs people hold about others' cooperativeness are likely shaped by everyday experiences of cooperativeness and trustworthiness in one's society. Generalizing from those four countries, a tentative conclusion is that economic and societal factors that influence beliefs (such as trust and trustworthiness) are likely key to explain societal variation in voluntary cooperation.

Second, our results provide an explanation of the key behavioral mechanisms that generate societal differences in voluntary cooperation, but because $a_i(b_i) = \widehat{c}_i \approx c_i$ in all subject pools, they also vindicate the ABC approach more generally. Therefore, the applicability of the ABC methodology extends beyond the cross-societal context in which we apply it here: the ABC methodology is a 'toolbox' that can be used to analyze the voluntary cooperation of any group of people, with or without the aim of comparing groups.

2. Related literature and hypotheses

Our paper builds on and contributes to a literature in behavioral economics that has documented some profound cross-societal differences in social preferences as observed in economic experiments (see, e.g., Gächter et al. (2010); Thöni (2019) for overviews)[4]. In this section, we review the literature most closely related to the behavioral mechanisms of voluntary cooperation. We are particularly interested in how existing societal differences in voluntary cooperation can be explained by the ABC approach, which can account for actual cooperation levels in one-shot games conducted in Western subject pools (e.g., Fischbacher et al. (2012); Gächter et al. (2017); Isler et al. (2021); Gächter et al. (2022)). We also review related literature on our second behavioral mechanism: punishment of free riding.

An important element of our first behavioral mechanism is *conditional cooperation*. Most experiments that have elicited conditionally cooperative attitudes have been conducted in WEIRD societies, that is, the US, UK and other Western societies (Thöni and Volk (2018))[5]. Exceptions include Herrmann and Thöni (2009); Martinsson et al. (2013) and Rustagi et al. (2010). Herrmann and Thöni (2009) conducted experiments in four Russian subject pools. Conditional cooperation was similar between them and also comparable to subject pools in Switzerland (Fischbacher et al. (2001)). Martinsson et al. (2013) elicited conditionally cooperative attitudes in

[4] For instance, Henrich et al. (2005) documented differences in ultimatum and dictator game behavior (in small-scale societies), see also Oosterbeek et al. (2004). Bohnet et al. (2008) and Bohnet et al. (2010) showed cross-societal differences in betrayal aversion; Chuah et al. (2016) and Chuah et al. (2023) observed cross-societal differences discrimination in a trust game environment; Romano et al. (2017) investigated ingroup and outgroup trust across 17 societies; Molleman and Gächter (2018) studied cross-societal differences in social learning and cooperation; and Gächter and Schulz (2016) and Cohn et al. (2019) found that honesty differs across societies. The biggest effort in establishing broad knowledge about how economic preferences (time, risk, and social) are distributed around the world is the Global Preference Survey (Falk et al. (2018)). For a seminal paper on cross-societal differences using observational data (on norms of corruption) see Fisman and Miguel (2007).

[5] Of the 18 studies considered in Thöni and Volk (2018), 16 were conducted in six WEIRD societies (USA, UK, Switzerland, Austria, The Netherlands, Denmark); 1 study included Japan (Kocher et al. (2008)); 1 study was done in Russia (Herrmann and Thöni (2009)). Studies not included in Thöni and Volk (2018) are, e.g., Kocher et al. (2015) (Germany); Bigoni et al. (2019) (Italy); Boosey et al. (2020) (USA); Bilancini et al. (2022) (mostly USA, UK, Ireland); Rustagi et al. (2010) (Ethiopia); Rustagi (2022) (Switzerland); and Martinsson et al. (2013) (Colombia and Vietnam).

Vietnam and Colombia, again with similar results to those from Western subject pools. Finally, Rustagi et al. (2010) elicited conditional cooperation in community forest groups in Ethiopia and found similar rates of conditional cooperation compared to those reported from Western societies (Thöni and Volk (2018)).

Beliefs about others' cooperativeness are a further element of our first behavioral mechanism. Beliefs have not been regularly studied in a cross-societal context. Beliefs are particularly interesting in a one-shot situation and among anonymous players who do not know one another because in this case beliefs can only be "home-grown", that is, influenced by daily economic, social, and political experiences outside the lab.

There is some evidence to suggest that we should expect higher levels of cooperative beliefs in our US and UK subject pools than in the Turkish and Moroccan samples. For example, data from the World Values Survey (WVS, Inglehart et al. (2014)) suggest that the proportion of people who believe that 'most people are trying to take advantage of me' is lower in the UK (5 %) and US (6 %) than in Turkey (14 %) and Morocco (20 %). More generally, across all countries in the WVS ($n = 58$), these beliefs are negatively correlated with an indicator of the rule of law (Spearman's rho = –0.38, $p = 0.003$) and an indicator of government effectiveness (rho = –0.44, $p = 0.001$). A similar picture emerges for 'generalized trust' ($n = 57$): here the correlation with the Rule of Law indicator is rho = 0.46, $p < 0.001$ and with the Government Effectiveness indicator rho = 0.52, $p < 0.001$. For further details, see Fig. A1 in Appendix A. Thus, together with behavioral evidence that trust – as a belief in the benevolent motives of others – has been shown to be positively correlated with cooperation in public goods games (Gächter et al., 2004; Thöni et al., 2012; Balliet and Van Lange, 2013; Kocher et al., 2015), we expect that beliefs about others' cooperation are higher in the UK and the US than in Morocco and Turkey.

Regarding our methodology to study our first behavioral mechanism, a related paper is Bigoni et al. (2019) who study a North-South divide in cooperation in Italy (Bigoni et al., 2016). They found that conditional cooperation was similar between participants from the North and the South. Beliefs about others' cooperation, however, were higher in the North than in the South. Despite its North-South divide, Italy is one country, with internal migration and a shared linguistic, cultural, and institutional heritage. Our societies arguably differ substantially more between each other than Italy differs internally.[6] Therefore, our study extends the evidence about the scope of cross-societal differences in the behavioral mechanisms underpinning voluntary cooperation by focusing on two countries that are culturally, economically, and institutionally among the most distant from the two prototypical WEIRD societies of the US and the UK.

In summary, evidence on conditionally cooperative attitudes predominantly comes from Western societies but results from the few non-Western studies are similar.[7] However, none of these studies conducted a systematic comparison of cooperative attitudes in Western societies and societies that are culturally distant from them within one comparable design, nor do they elicit beliefs about others' cooperativeness and the willingness to punish other group members. Providing a fully comparable analysis is thus one contribution of our paper.

Based on existing literature, we formulate the following hypotheses: we expect little difference in conditionally cooperative attitudes between our four subject pools (Hypothesis 1a), but higher cooperative beliefs in the UK and US subject pools (Hypothesis 1b), which will lead to higher cooperation in those countries (Hypothesis 1c).

The second behavioral mechanism concerns *punishment*. Closest to us on this dimension are Herrmann et al. (2008) and Gächter and Herrmann (2009) who also studied cooperation and punishment, but did not elicit cooperative attitudes or beliefs. Herrmann et al. (2008) used a repeated game in fifteen (western and non-western) countries around the world. They found little differences in how people punish free riders, but strong differences in "antisocial punishment", that is, how people punish those who contribute the same or more than them. Antisocial punishment was largely absent in their UK and US subject pools but was substantial in their Turkish and Arabic subject pools (see also Bruhin et al. (2020) for a further analysis of this result). Likely explanations include exposure to persistent resource scarcity (Prediger et al., 2014) or strategic revenge (e.g., Sylwester et al. (2013)), which was possible given the repeated game design. Our one-shot design excludes this possibility.

If antisocial punishment is not only an act of revenge in a repeated interaction, but also a "preference", it might be observed even in our one-shot design. The one-shot experiments of Gächter and Herrmann (2011) in Russia, and Gächter and Herrmann (2009) in Russia and Switzerland, suggest this possibility. Antisocial punishment did occur in their Russian samples but was negligible in their Swiss subject pools. If these results generalize, then we might find no subject-pool differences in the punishment of free riders, and more antisocial punishment in Morocco and Turkey than in the US and the UK. This is our Hypothesis 2a.

Regarding cooperation in the presence of punishment, the results from Gächter and Herrmann (2009) showed that cooperation was higher in Switzerland than in Russia. Similarly, Herrmann et al. (2008) found that cooperation was higher in subject pools with lower antisocial punishment. We therefore expect that cooperation in the presence of punishment will be higher in the UK and US than in Moroccan and Turkish subject pools: our Hypothesis 2b.

[6] For other within-country (culture) experiments measuring social preferences, see, e.g., Brosig-Koch et al. (2011), Kim et al. (2017) and Choi et al. (2020).

[7] Comparing the rate of cooperative attitudes across the 20 WEIRD and 5 non-WEIRD samples using data from Rustagi et al. (2010), Martinsson et al. (2013), and Thöni and Volk (2018) shows a similar percentage of conditional cooperators ($M_{\text{WEIRD}} = 59$ %; $M_{\text{non-WEIRD}} = 51$ %; Mann-Whitney $Z = 1.50$, $p = 0.148$) and a significantly higher percentage of free riders in WEIRD societies ($M_{\text{WEIRD}} = 21$ %; $M_{\text{non-WEIRD}} = 12$ %; Mann-Whitney $Z = 1.97$, $p = 0.048$).

3. Conceptual background and methods

3.1. Conceptual background of the behavioral mechanisms of voluntary cooperation

Our first behavioral mechanism, the ABC approach, builds on Fischbacher and Gächter (2010); Fischbacher et al. (2012); Gächter et al. (2017); and in particular Isler et al. (2021). Isler et al. introduced a version of the ABC approach according to which contextual features (framing in their case) can affect attitudes a_i and beliefs b_i, but together a_i and b_i predict c_i equally well for all contextual features. In our case the relevant context is societal background: Individual i who lives in society s decides their contribution c_i^s as a function of their cooperative attitude a_i^s and beliefs b_i^s about others' average contribution: $a_i^s(b_i^s) \rightarrow c_i^s$. This implies that voluntary cooperation levels c_i^s might differ between societies because a_i^s or b_i^s differ between societies. In this paper we will show that societal variation in c_i^s are largely driven by variation in the beliefs, b_i^s.

Regarding punishment, i's preferences for punishment of another individual j, $p_{ij}^s(c_i^s, c_j^s)$, depends on i's and j's contributions and might be society (group/culture) specific (Herrmann et al. (2008); Gächter et al. (2010)). Ultimately, we thus expect voluntary cooperation in an environment with peer-punishment to depend on attitudes, beliefs, and expected punishment: $c_i^s = f(a_i^s(b_i^s), e(p_{ji}^s))$.[8]

Long-run historical processes, such as whether society s tends towards an "individualistic" or "collectivistic" culture (e.g., Greif (1994)), can shape the psychological dispositions a_i^s and p_{ij}^s and beliefs b_i^s by which people approach cooperation. Historical processes also shape the institutions around which society s organizes its economic and social life (Schulz et al., 2019; Schulz, 2022; Henrich, 2020; Rustagi, 2022). Beliefs b_i^s play a central role because actual experiences with institutions, combined with processes of cultural and parental transmission of values (e.g., Boyd and Richerson (1985); Boyd and Richerson (2009); Bisin and Verdier (2011); Hauk and Saez-Marti (2002); Guiso et al. (2008); Tabellini (2008); Henrich (2015)), are likely to shape beliefs b_i^s. While we do not investigate the causal role of historical, cultural, or institutional factors on cooperative attitudes and beliefs, by comparing societies that vary substantially along cultural dimensions we contribute to the understanding of the relative importance of the two underlying factors in explaining differences in voluntary cooperation.

Our one-shot experiments are designed to provide measures of the key variables of interest while excluding any strategic incentives that might be present in repeated games. These are: cooperative attitudes a_i^s, beliefs b_i^s, contributions c_i^s, punishment p_{ij}^s, and the expected punishment of i by j, $e(p_{ji}^s)$, from four societies s: Morocco, Turkey, UK, and US. Details are in subsection 3.2.

All subjects participate in two games, the N-Game followed by the P-Game. The N-Game measures the first behavioral mechanism $a_i^s(b_i^s) \rightarrow c_i^s$ in the absence of punishment. The P-Game provides the data for our second behavioral mechanism, which measures $c_i^s = f(a_i^s(b_i^s), e(p_{ji}^s))$. In the next section we describe the rationale for our selection of countries and the evidence about their cultural distance. Details are in Sections 3.4 and 3.5.

3.2. The cultural, economic, and institutional distance between the four subject pools

The UK and US share several cultural, economic, and institutional characteristics. These are substantially different from those of Morocco and Turkey, which also share important similarities with each other. Here, we describe the most important ones; the details are in Appendix A and illustrated in Figs. A2 and A3.

Regarding cultural differences, a frequently-used cultural dimension is individualism (e.g., Hofstede and Hofstede (2001)). The prototypically WEIRD countries US and the UK are among the most individualistic societies in the world whereas the non-WEIRD countries Morocco and Turkey are more collectivistic. According to Inglehart and Welzel (2005), the UK and the US have similarly high "self-expression values", whereas Morocco and Turkey, share more "survival values". See Schulz et al. (2019) and Henrich (2020) for recent investigations into the psychology underpinning WEIRD cultures and how they differ from non-WEIRD ones.

To establish the cultural distance between our subject pools as systematically as possible, we rely on the most comprehensive dataset available (Muthukrishna et al., 2020). Using data on approximately 170,000 individuals from 80 countries, representing approximately 85 % of the world's population, Muthukrishna et al. (2020) developed a quantitative measure of cultural distance between these 80 societies. Cultural distance is calculated as the ratio of between- and within-country variance in bilateral country comparisons of individual responses to questions in the World Values Survey (Waves 2005–09 and 2010–14). The questions cover a wide range of values and beliefs (e.g., attitudes to private matters, family and religious values, attitudes to work and political values).

The quantitative measures of *cultural distance* between any pair of countries provide further evidence for the two distinct clusters of the societies we selected. In terms of ranked cultural distance using the US as the baseline, the UK is the 6th closest country to the US; Turkey the 47th; and Morocco the 59th (Egypt is ranked 80th, that is, farthest away from the US). Taking Morocco as the baseline shows that Turkey is ranked 18th; the US is ranked 60th, and the UK is ranked 71st (farthest away is Sweden, ranked 80th).

Our subject pools also differ starkly in their *economic success*. For instance, using the GDP per capita at purchasing power parity as an indicator shows that the US and the UK are among the richest countries in the world, whereas Turkey and Morocco are substantially poorer. *Institutional quality*—as measured by the countries' Government Effectiveness and Rule of Law indicators of the Worldwide Governance Indicators (Kaufmann and Kraay, 2016)—, is among the highest in the world in the US and UK, and substantially lower

[8] Based on the results of Weber et al. (2018) who found that conditionally cooperative attitudes and likelihood to punish are unrelated, we expect that one's preference for punishment (p_{ij}^s) is independent of one's cooperative attitude (a_i^s).

(around the world average) in Turkey and Morocco.

From this diverse set of indicators we conclude that our subject pools come from two culturally, economically and institutionally distant clusters, but share similarities within their respective cluster. The extent to which these *ex ante* similarities and differences in the societies of our subject pools are relevant *ex post* for the behavior of our subject pools is the focus of our study. If we find differences in conditional cooperation, beliefs, and punishment between our subject pools, we expect them to be greater between clusters than within them. Note that the assumption of two cultural clusters is not strictly necessary to test the behavioral mechanisms described above, as it is potentially interesting to compare any two countries. However, there is an important advantage of doing so: If differences between clusters are larger than differences within clusters, then these are likely driven by cultural/societal distance, provided that other subject pool differences are controlled for.

3.3. Participants and procedures

We follow Herrmann et al. (2008) in choosing student samples with comparable age and socio-economic status within their respective society because societies likely differ in their socio-economic profiles. Choosing student samples ensures subject pool comparability (in terms of education levels, and age, which has been shown to positively influence cooperation (see, e.g., List (2004); Gächter and Herrmann (2011); Arechar et al. (2018)). Comparability minimizes confounds that come from the differing socio-demographic composition of the respective subject pools (see also Gächter (2010); Thöni (2019) for conceptual discussions of this point). We measure additional (and unavoidable) variation in our post-experimental questionnaire and use them as controls in our regression analyses. Our objective is more methodological and not to provide representative country-wide measures of cooperation (these would require sample sizes of about 1000 people in each country and would be infeasible for controlled lab experiments such as ours) but rather to document existing differences in cooperation (among *comparable* subpopulations drawn from different countries) and shed light on the behavioral mechanisms that give rise to those differences.

A total of 388 students participated in our study; 93.8 % of participants were nationals of their respective country.[9] We recruited 128 students at Stony Brook University in the US; 92 students at the University of Nottingham in the UK; 80 students at the École Nationale d'Agriculture de Meknès in Morocco; and 88 students at Istanbul Bilgi University in Turkey. Table 1 summarizes the characteristics of our samples.

The experiments were computerized and conducted with z-Tree (Fischbacher, 2007). At the École Nationale d'Agriculture de Meknès the experimenters recruited participants from the local student population. At the other three universities, we used the recruitment software ORSEE (Greiner, 2015) to invite participants. In all samples, subjects could only participate once in the experiment.

We follow the established rules for conducting cross-societal economic experiments (Roth et al., 1991; Thöni, 2019). In all four laboratories, the experiments were run by local research assistants. The instructions and software used the same experimental currency units across all laboratories and were presented in the local language.[10] Participants made their decisions in private with visual separations between workstations. The sessions lasted approximately 90 minutes and were conducted according to a strict protocol to minimize the differences in the way sessions were run across countries. We paid participants in private at the end of each session. Experimental currency was transferred into the local currency using different exchange rates. The exchange rates were chosen to reflect local purchasing power and student wages so that, in real terms, the possible real earnings were similar across subject pools. The average payoffs in local currency are reported in Table 1.

Each session consisted of two experimental games (the N-Game followed by the P-Game) and a socio-economic questionnaire. Participants were randomly re-matched after the N-Game. We did not provide any feedback after the N-Game to prevent participants from updating their beliefs and to reduce potential income effects and strategic play in the P-Game. However, as will become clear in Section 3.5, in contrast to the N-Game, the P-Game requires feedback on others' contribution, which is why the P-Game always followed the N-Game.

3.4. Public good game with no punishment (N-Game): ABC elicitation

To measure cooperative attitudes, we used the strategy method for public goods games, developed by Fischbacher et al. (2001). Participants were randomly allocated to groups of four. Each received an endowment of 20 tokens and decided how much to contribute to a common "project". Individual payoffs were determined by the following function:

$$\pi_i = 20 - g_i + 0.4 \cdot \sum_{j=1}^{4} g_j. \tag{1}$$

Participant i's contribution to the public good is given by g_i, and the size of the public good is the sum of all contributions. The

[9] We excluded 28 participants from the sample of 388 students (22 participants from the US, 4 participants from the UK and 2 participants from Turkey) who indicated in the post-experimental questionnaire that they were not citizens of the respective countries. All main results of this paper hold independently of their inclusion.

[10] The instructions are available in Online Appendices J and K. We translated the instructions from English into Arabic or Turkish and had them back translated by another person in order to make sure that instructions were as identical as possible (Brislin (1970)).

Table 1
Characteristics of the four samples.

	US	UK	Morocco	Turkey
Average age (*SD*)	19.56	19.63	20.84	22.02
	(2.75)	(1.85)	(1.44)	(1.93)
Female	59 %	56 %	65 %	40 %
Business or economics	9 %	23 %	1 %	22 %
Urban background	46 %	63 %	48 %	67 %
Average number of siblings	1.69	1.69	3.03	1.79
(*SD*)	(1.35)	(1.37)	(2.18)	(2.01)
Middle class	58 %	84 %	89 %	42 %
No. obs.	106	88	80	86
Average payoff in local currency (*SD*)	USD 13.87	GBP 11.50	MAD 100.46	TRY 40.93
	(3.32)	(1.53)	(21.64)	(7.95)

Notes: Business and Economics: percentage of participants studying business or economics. Urban background: percentage of participants who lived most of their life in a town with at least 10,000 inhabitants. Middle class: percentage of participants who self-reported their family income at age sixteen to be at least average compared to other families.

marginal per capita return of the public good is 0.4 monetary units (MU). Although the social optimum entails full contributions, the individually money-maximizing strategy is to contribute nothing.

Participants made two decisions in the N-Game: they chose an *unconditional contribution* c_i to the project and filled in a *contribution table*. This table allowed for conditioning the participant's own contribution on all rounded average contributions $\overline{c_j} \in \{0, 1, 2, \ldots, 20\}$ of her three other group members j and provides us with our measurement of the function $a_i(\overline{c_j})$. To ensure incentive compatibility, the actual contribution for one randomly chosen participant per group was taken from their contribution table according to the average contributions of the three co-players.

After making contribution decisions, we elicited participants' *beliefs* about the other group members' average unconditional contribution. Participants earned three MUs for guessing correctly, two MUs for a deviation of one point, one MU for a deviation of two points and zero MUs for a higher deviation (Gächter and Renner, 2010).[11] We did not provide feedback after the N-Game and fixed this sequence across all sessions to prevent participants' from updating their beliefs about other participants' behavior.

Using the criteria outlined in Thöni and Volk (2018), we classify participants into different types of cooperative attitudes according to their contribution tables. *Conditional Cooperators* (CC) show a positive correlation between their own conditional contributions and the average contributions of their fellow group members (i.e., Pearson's $\rho \geq 0.5$) or at least one increase in their contribution schedule. *Free Riders* (FR) contribute nothing for every possible average contribution of their group members. We refer to participants who are not classified as either CC or FR as *Unclassified Others* (OT). The results from the N-Game provide the data for our first behavioral mechanism $a_i^s(b_i^s) \rightarrow c_i^s$ presented in Section 4.

3.5. Public good game with punishment (P-Game): Punishment elicitation

The P-Game consists of two stages: First, participants choose their contribution in a one-shot public goods game for which the individual payoff function is given by (1). Then participants state their beliefs about the other group members' average contribution to the public good. We did not incentivize beliefs in the P-Game to avoid punishment motivated by disappointment due to wrong beliefs or income effects (Cubitt et al., 2011). In the second stage of the game, participants learned their individual contributions of their group members and could assign up to five punishment points to each. Each assigned punishment point cost the punisher one MU and removed two MUs of the targeted person's income. Individual payoffs were determined by the following function:

$$\pi_i = \pi_i^{S1} - \sum_{j=1}^{4} p_{ij} - 2 \cdot \sum_{j=1}^{4} p_{ji}. \tag{2}$$

Participant i's payoff from the first stage is given by π_i^{S1}. The punishment points group member i allocates to group member j are denoted by p_{ij}, and p_{ji} denotes the punishment points allocated by j to i. Each participant received 10 MUs to cover potential losses. In the P-Game, the punishment $e(p_{ji}^s)$ that group member i expects to receive might influence i's contribution.

After assigning punishment points to their group members, participants were asked to indicate how many punishment points they expected to receive from each of their group members. The results from the P-Game provide the data for our second behavioral mechanism $c_i^s = f(a_i^s(b_i^s), e(p_{ji}^s))$ presented in Section 5.

[11] Belief hedging might be a concern in this design. However, we think this is unlikely for two reasons. First, incentives to hedge are minor here. Second, Blanco et al. (2010) find no evidence for belief hedging in a similar social dilemma experiment.

4. Behavioral mechanism I: attitudes and beliefs

4.1. Preferences for conditional cooperation ("attitudes")

We first classify the participants according to their individual cooperative attitudes (see Section 3.4) and compare the distributions across societies in Fig. 1a (a_i^s in our behavioral mechanism).[12] Using χ^2 tests, we find that the distributions of cooperative attitudes are statistically significantly different across the four samples ($p = 0.043$). The most common cooperative attitude in all four samples is conditional cooperation (CC), with a statistically similar share of CC across the four subject pools ($\chi^2(3) = 1.70, p = 0.637$). However, the shares of free riders (FR; $\chi^2(3) = 9.19, p = 0.027$) and the shares of others (OT; $\chi^2(3) = 6.67, p = 0.083$) vary somewhat across the four samples.[13]

Comparing the UU cluster (pooling the US and UK samples) to the MT cluster (pooling the Morocco and Turkey samples) reveals statistically similar distributions of cooperative attitudes ($p = 0.220$). Within clusters, there are no significant differences between the cooperative attitudes of the US and UK ($p = 0.123$) samples and weakly significant differences between the Morocco (MA) and Turkey (TR) samples ($p = 0.053$).

Another way of investigating sample differences in cooperative attitudes is to compare the *extent* of conditional cooperativeness. We use the average contributions in the contribution table of the N-Game as a measure of conditional cooperativeness (Fig. 1b). Across the four samples, on average participants increase their conditional contributions as the average contribution of others rises, but they only imperfectly match the average contribution of others by undercutting others' contribution. Imperfect conditional cooperation results in average conditional contributions below the 45°-line (which corresponds to perfect conditional cooperation). Fig. 1b also shows that sample differences in average conditional contributions are small in the "behaviorally relevant" interval between 0 and 11 where 91 % of the average unconditional contributions of other group members fall. Fig. 1c depicts average conditional cooperation for conditional cooperators (CC) only, and Fig. 1d for others (OT) only.

Finally, we use a regression analysis to test for across-cluster and within-cluster differences in the level and slope of the participants' contribution schedules (Table 2).[14] Again, we refer to the UK and the US as the UU cluster and to Morocco and Turkey as the MT cluster. First, we investigate across-cluster differences in a pooled estimation by regressing the conditional contribution on the average contribution of other group members, a dummy variable for the MT cluster and an interaction term between the dummy and the average contribution of others to measure differences in slopes (Table 2, Col. 1). Additionally, we control for the socio-economic background of participants.

The coefficient for the average contributions of others is positive and highly significant, but smaller than 1 implying imperfect conditional cooperation in the UU cluster. Conditional contributions increase on average by 0.438 for each additional money unit contributed. The negative and highly significant interaction term implies a smaller increase in conditional contributions in the MT cluster. The statistically insignificant MT cluster dummy indicates a similar average conditional contribution in both clusters.[15]

Next, we estimate a similar model with country dummies and country-specific interaction terms to test for within-cluster variation (Table 2, Col. 2). Here, the US sample serves as the comparison group. The statistical insignificance of the UK dummy and its interaction term "Average contribution of others × UK" imply no variation in the slope or levels within the UU cluster comparing the US and UK samples. Additionally, we find no statistically significant differences in the coefficients of "Average contribution of others × MA" and "Average contribution of others × TR" ($F(1, 359) = 2.16, p = 0.142$) or the MA and TR dummies ($F(1, 359) = 0.03, p = 0.858$). This indicates similar slopes and levels of conditional contributions within the MT cluster.

We summarize our findings on cooperative attitudes as follows:

Result 1a: *We find a similar distribution of cooperative attitudes across our four subject pools, but a slightly higher conditional cooperativeness in the UU compared to the MT cluster.*

Result 1a provides mixed support for Hypothesis 1a, which stipulates no differences between subject pools. While the distribution of cooperative attitudes is similar across subject pools, we do find slightly weaker conditional cooperativeness in Morocco and Turkey than in the UK and US. Our next step is to test Hypothesis 1b that beliefs about others' cooperation are higher in the UK and US than in Morocco and Turkey.

4.2. Beliefs and contributions

Fig. 2a shows *beliefs* about unconditional contributions in the N-Game (b_i^s in the ABC framework). We find significant variation in beliefs across the four samples (Kruskal-Wallis test, $\chi^2(3) = 18.58$, $p < 0.001$). Beliefs are significantly higher in the UU cluster

[12] The results reported here are robust to an alternative classification (see Fig. 1B in Online Appendix B).

[13] CC include perfect conditional cooperators, who exactly match others' contribution (13 in the US, 9 in the UK, 3 in MA, and 3 in TR). OT include unconditional cooperators (3 in the US, 2 in the UK, 0 in MA, 9 in TR), triangle contributors (24 in the US, 16 in the UK, 23 in MA, 6 in TR) and unclassified others (9 in the US, 5 in the UK, 13 in MA, 14 in TR).

[14] When estimating the same regression models including only CC, we find slight variations in conditional contributions across the two clusters and within the MT cluster (Table C1 in the Online Appendix C). Using the same method to compare the conditional contributions of OT shows only significant level differences within the UU cluster and, overall, a very poor fit of the regression model (Table C2 in the Online Appendix C).

[15] The estimation results for the UK (Model (2) in Table 3) are similar to those from Weber et al. (2018) who also ran a comparable strategy method public goods game with UK subjects. In their data, the coefficient (*SE*) for average contribution of others is 0.401 (0.022).

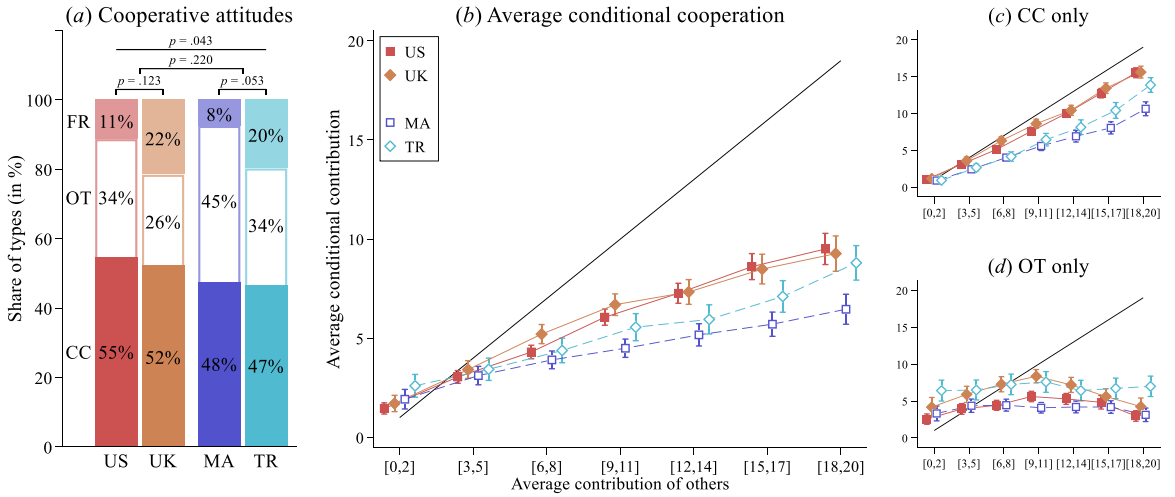

Fig. 1. Cooperative attitudes (a_i^s in our behavioral mechanism). Panel (*a*): The share of conditional cooperators (CC, dark bars), free riders (FR, white bars) and others (OT, light bars); percentages indicate frequency of types; *p*-values from (pooled) χ^2 tests. Panel (*b*): Average contribution schedule by country in the N-Game; all participants. Panel (*c*): Conditional cooperators (CC) only. Panel (*d*): Others (OT) only. The 45°-line corresponds to perfect conditional cooperation. US = United States; UK = United Kingdom; MA = Morocco; TR = Turkey. The error bars indicate ± 1 *SEM*.

Table 2
Conditional contributions across subject pools.

Dependent variable: conditional contribution	(1) By cluster	(2) By country
Average contributions of others	0.438***	0.456***
	(0.034)	(0.044)
Average contribution of others × MT cluster	−0.152***	
	(0.045)	
MT cluster	0.653	
	(0.509)	
Average contribution of others × UK		−0.040
		(0.068)
UK		0.851
		(0.588)
Average contribution of others × MA		−0.216***
		(0.060)
MA		1.021*
		(0.593)
Average contribution of others × TR		−0.127**
		(0.063)
TR		1.171
		(0.755)
Socio-economic controls	Yes	Yes
Constant	3.016	3.062
	(2.655)	(2.710)
R^2	0.14	0.15
No. obs. (Clusters)	7560	7560
	(360)	(360)

Notes: OLS estimation with robust *SE* clustered on individuals in parentheses. The dependent variable is the conditional contribution elicited on the contribution table in the N-Game where participants indicated their own conditional contribution for each of the 21 average contributions of their group members. UU cluster: UK and US; MT cluster: Morocco and Turkey. Socio-economic control variables: age, female, urban background, middle class, single child, economics/business student. None is significant at any conventional level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

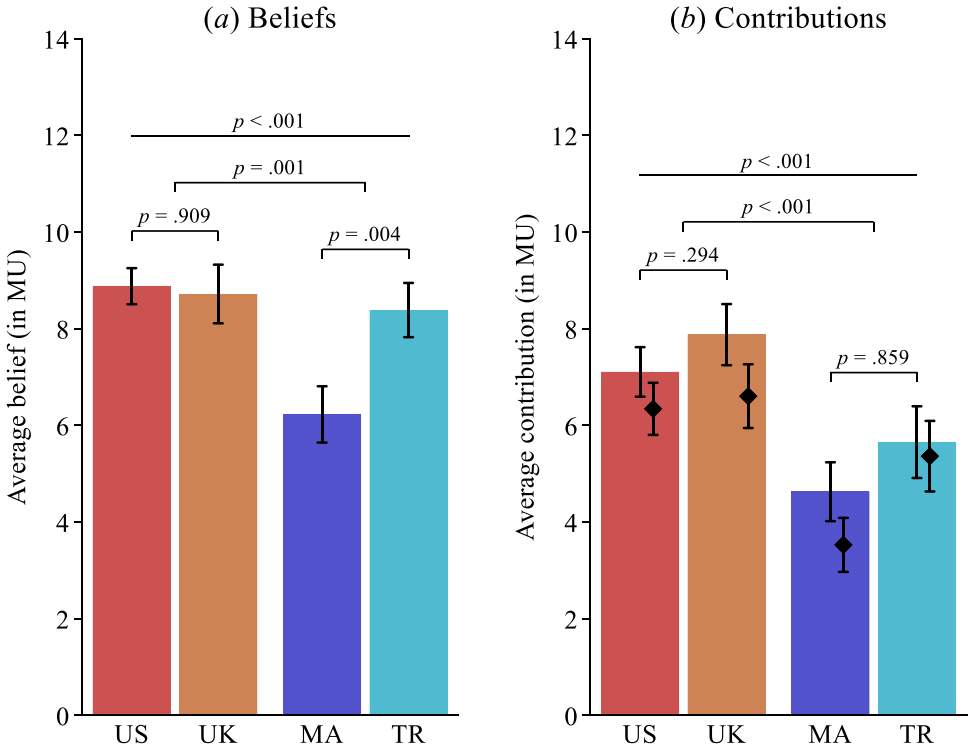

Fig. 2. Average beliefs (Panel *a*); and average unconditional contributions to the public good (Panel *b*) in the N-Game by country. Diamonds (◆) indicate the average predicted contributions based on cooperative attitudes and beliefs (described in section 4.3). US = United States; UK = United Kingdom; MA = Morocco; TR = Turkey. The error bars indicate ±1 *SEM;* *p*-values from Kruskal-Wallis and (pooled) Mann-Whitney tests.

compared to the MT cluster. We do not find significant differences in beliefs within the UU cluster. For the MT cluster, we find highly significant differences between the MA and TR samples. The between-cluster difference in beliefs is driven by lower beliefs in the MA sample; beliefs in the TR sample are not significantly different from those held in the US or UK sample (Mann-Whitney $Z = 1.11$, $p = 0.269$; $Z = 0.52$, $p = 0.603$; resp.).[16]

We collect these findings in Result 1b, which confirms our Hypothesis 1b:

Result 1b. *Beliefs about others' cooperation are higher in the UU cluster compared to the MT cluster.*

We now turn to Hypothesis 1c which predicts that cooperation is higher in the UU cluster than in the MT cluster. Fig. 2b illustrates the average unconditional *contributions* in the N-Game (c_i^s in the ABC framework) which vary significantly across the four samples (Kruskal-Wallis test, $\chi^2(3) = 19.62$, $p < 0.001$). Unconditional contributions in the UU cluster are significantly higher than in the MT cluster. There are no differences in unconditional contributions within-clusters when comparing the US and UK samples or the MA and TR samples. The next result summarizes this finding:

Result 1c *Unconditional contributions are higher in the UU than in the MT cluster.*

This supports Hypothesis 1c. Our next step is to explain contributions as a function of cooperative attitudes and beliefs.

4.3. Explaining contributions as a function of cooperative attitudes and beliefs: $a_i^s(b_i^s) \rightarrow c_i^s$

To explain contributions, we calculate the variable "Predicted contribution" by combining a participant's individual contribution table and their belief about others' contributions ("Belief") as elicited in the N-Game. Formally put, we compute for each individual i their predicted contribution $\widehat{c}_i$ as $a_i(b_i) = \widehat{c}_i$. The sample average predicted contributions $\widehat{c}^s$ are shown as diamonds (◆) in Fig. 2b.

The predicted contributions vary significantly across the four samples (Kruskal-Wallis test, $\chi^2(3) = 15.03$, $p = 0.002$) and tend to be lower than the actual contributions (Fig. 2b). Predicted contributions in the UU cluster are significantly higher than in the MT cluster (pooled Mann-Whitney $Z = 3.73$, $p < 0.001$). There are no differences in predicted contributions when comparing the US and UK samples (Mann-Whitney $Z = -0.09$, $p = 0.931$) and the MA and TR samples (Mann-Whitney $Z = -0.65$, $p = 0.517$). However, the predicted contributions are significantly lower than the actual contributions for all samples except Turkey (Wilcoxon signed-rank test $Z_{US} = 2.06$, $p_{US} = 0.040$; $Z_{UK} = 2.88$, $p_{UK} = 0.004$; $Z_{MA} = 1.52$, $p_{MA} = 0.131$; $Z_{TR} = 1.89$, $p_{TR} = 0.059$).

[16] We also test for differences in the accuracy of beliefs—defined as deviation from the actual contribution of others—and find a statistically similar belief accuracy across samples (overall average 1.8, Kruskal-Wallis test, $\chi^2(3) = 4.31$, $p = 0.230$; for details see Table D1 in the Online Appendix D).

We also test for the accuracy of predicted contributions ($\widehat{c_i^s}$) across samples. Comparing predicted with actual contributions in the N-Game yields accuracy ($\widehat{c_i^s} = c_i^s$) of over 61 % for all four societies and no significant differences in the accuracy of predicted contributions across subject pools (Fig. E2 in the Online Appendix E).[17]

To explain the gap between actual and predicted contributions, we follow Fischbacher and Gächter (2010) (see their Table 2) and regress unconditional contributions (c_i^s) on "Predicted contribution" ($\widehat{c_i^s}$), "Beliefs" (b_i^s), a MT cluster dummy and interaction terms in a pooled model (Table 3, Col. 1). We also control for the socio-economic background of participants. "Belief" is included to test for cooperation above $\widehat{c_i^s}$—as observed by Fischbacher and Gächter (2010).

The coefficients "Predicted contribution" and "Belief" are positive and highly significant, confirming an earlier result by Fischbacher and Gächter (2010). This suggests that participants—most of whom are imperfect conditional cooperators with contribution schedules below the 45-degree line—are more cooperative than predicted by their preferences ($\widehat{c_i^s}$) alone. The results show no significant differences in the level of contributions across clusters (MT cluster: $b = 0.164$, $p = 0.850$) when controlling for beliefs, suggesting that the observed differences in behavior may be driven by variation in beliefs. There are also no differences in the way "Predicted contributions" and "Belief" are associated with contributions across clusters. This follows from the insignificant dummy variable "MT cluster", and the interaction terms "Predicted contribution × MT cluster" and "Belief × MT cluster".

The significant influence of beliefs on contributions across both clusters shows that conditional cooperation is greater than the level implied by "Predicted contribution" alone, that is, "Belief" matters in addition to "Predicted contribution", a finding also reported by Fischbacher and Gächter (2010).[18]

Next, we turn to the within-cluster comparisons (Table 3, Col. 2). Within the UU cluster, we find neither level differences in contributions between the US and UK samples, nor significant differences in the way "Predicted contribution" and "Belief" are associated with the contributions. To investigate differences in the MT cluster, we test for significant differences in the coefficients of the MA and TR dummies, as well as the interaction terms with predicted contributions and beliefs. We find significant level differences in contributions ($F(1, 342) = 7.39$, $p = 0.007$) and a significant difference in the way beliefs are associated with contributions ($F(1, 342) = 6.79$, $p = 0.010$). However, we find no differences in the way "Predicted contributions" influence contributions across the MA and TR samples ($F(1, 342) = 0.51$, $p = 0.474$). This shows considerable within-cluster variation for the MT cluster. Contributions are higher in the MA sample compared to the TR sample, but participants from the TR sample show a greater increase in their contributions for a higher expected average contribution from others.

The results of the first regression model have two important implications: First, when we control for beliefs, we find no significant level differences in contributions between the two clusters (the MT cluster dummy is insignificant). Since "Belief" is highly significant, this suggests that differences in beliefs are an important driver of differences in cooperation between the subject pools. Second, the insignificant interactions between "Predicted contributions" and the cluster dummy indicate that a one-unit increase in "Predicted contributions" increases actual contributions by the same amount across clusters. The same interpretation applies to the non-significant interaction between "Belief" and the cluster dummy. This shows that the relative importance of "predicted contributions" and "Belief" in explaining behavior is similar across clusters. In Online Appendix I we report a simulation analysis that confirms the relative importance of beliefs in explaining differences in contributions across clusters: 37 % of the difference is due to beliefs, while only 14 % is due to cooperative attitudes.

4.4. Discussion

Regarding our hypotheses on the first behavioral mechanism, we find mixed support for Hypothesis 1a, which predicts similar distributions of conditionally cooperative attitudes a_i^s across the four subject pools. The distribution of conditionally cooperative attitudes, which have previously been elicited mainly in US and Western European subject pools (Thöni and Volk, 2018), generalizes to all our subject pools. However, the mean slopes are slightly lower in the MT cluster than in the UU cluster. Beliefs b_i^s about others' cooperation are higher in the UU cluster than in the MT cluster, confirming Hypothesis 1b. Beliefs appear to be an important factor explaining contributions in both clusters, including higher contributions in the UU cluster than in the MT cluster, confirming Hypothesis 1c. In summary, with respect to our first behavioral mechanism, the ABC approach ($a_i^s(b_i^s) \rightarrow c_i^s$), cooperative attitudes and beliefs explain unconditional contributions in all subject pools.

5. Behavioral mechanism II: punishment

Our next step is to investigate the second behavioral mechanism—expected ($e(p_{ji}^s)$) and actual punishment (p_{ij}^s)—and their

[17] The difference between predicted and actual contributions becomes smaller when looking at conditional cooperators only (Fig. E1 in the Online Appendix E). This observation is consistent with findings by Fischbacher et al. (2012).

[18] Reassuringly, the estimation results for the UK (Col. 1 in Table 3) are similar to those from Weber et al. (2018) who also ran a comparable strategy method public goods game with UK subjects. The coefficient (*SE*) for predicted contribution is 0.532 (0.067) and for beliefs it is 0.392 (0.068).

Table 3
Explaining unconditional contributions in the N-Game.

Dependent variable: unconditional contribution	(1) By cluster	(2) By country
Predicted contribution	0.504***	0.463***
	(0.090)	(0.112)
Belief	0.300***	0.397***
	(0.116)	(0.141)
Predicted contribution × MT cluster	0.088	
	(0.150)	
Belief × MT cluster	−0.191	
	(0.159)	
MT cluster	0.164	
	(0.863)	
Predicted contribution × UK		0.049
		(0.181)
Belief × UK		−0.127
		(0.221)
UK		1.616
		(1.061)
Predicted contribution × MA		0.225
		(0.170)
Belief × MA		−0.536***
		(0.174)
MA		2.652**
		(1.201)
Predicted contribution × TR		0.081
		(0.190)
Belief × TR		−0.059
		(0.210)
TR		−0.903
		(1.144)
Socio-economic controls	Yes	Yes
Constant	2.042	1.026
	(2.390)	(2.369)
R^2	0.49	0.50
No. obs.	360	360

Notes: OLS estimates with robust *SE* in parentheses. UU cluster: UK and US; MT cluster: Morocco and Turkey. Socio-economic control variables: age, female, urban background, middle class, single child, economics/business student. None is significant at any conventional level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

implication for cooperation and beliefs ($c_i^s = f(a_i^s(b_i^s), e(p_{ji}^s))$).[19] For this purpose, we analyze the P-Game described in Section 3.5 and test Hypotheses 2a and 2b.

5.1. Expected punishment

After participants made their contribution decisions in the P-game and were informed about the contributions of their group members, we elicited expected punishment by asking participants to indicate the number of punishment points they expected to receive from each of their group members. Using the second behavioral mechanism described in Section 3.1, we analyzed expected punishment separately for negative and non-negative deviations from the potential punisher's contribution (Fig. 3). Expected prosocial punishment (that is, punishment for negative deviations from the punisher's contribution) differs significantly across the four subject pools. We find highly significant differences in expected prosocial punishment between the UU and MT clusters, with significantly lower expected punishment for free riding in the MA sample than in the other three samples.

Expected antisocial punishment (punishment of *non-negative* deviations from the punisher's contribution) is statistically similar across the four samples. It is also statistically similar when comparing the UU and MT clusters, as well as when comparing the US and UK samples or the MA and TR samples. Tobit regressions that control for the size of the negative or positive contribution deviations, and the socio-economic background of participants, largely confirm the results of our non-parametric analysis (Online Appendix G, Table G1).[20]

[19] In Online Appendix F, we also test—and confirm—the assumption of our behavioral mechanisms that p_{ij}^s is independent of a_i^s—see Section 3.1; and Weber et al. (2018); Molleman et al. (2019) who provide evidence for this assumption.

[20] A further question is whether the accuracy of expected punishment differs across samples. We find a significantly lower accuracy of expected punishment in the UU cluster compared to the MT cluster, due to overestimation of the use of punishment in the US and UK samples. We also find a significant within-cluster difference in accuracy for the MT cluster only (Fig. H1 in the Online Appendix H).

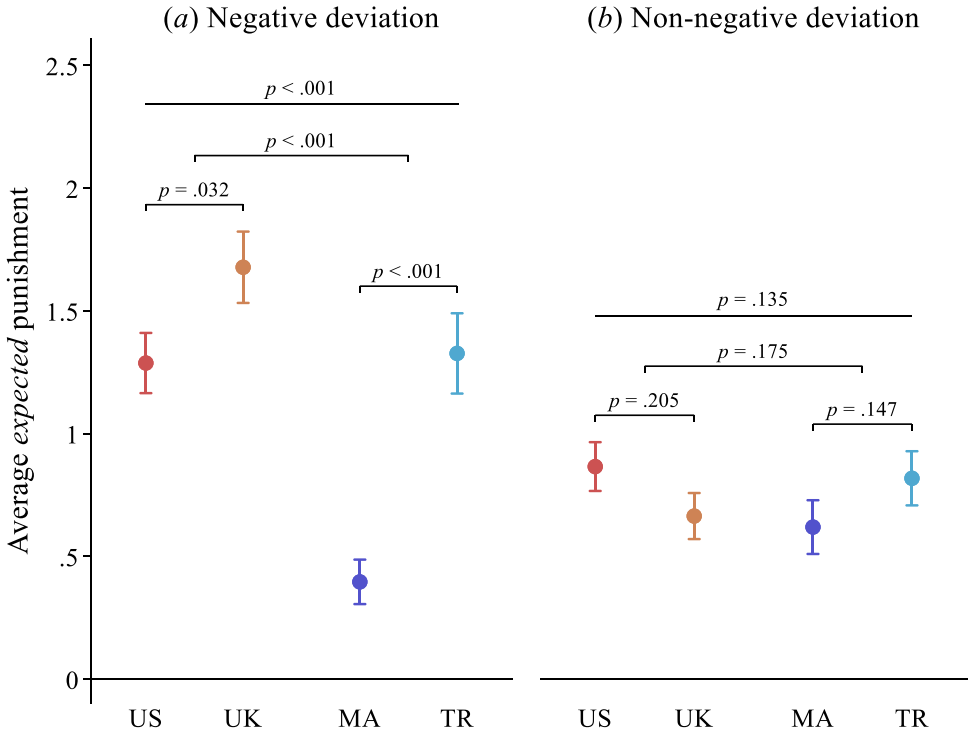

Fig. 3. Average *expected* punishment points by country for a negative deviation (Panel *a*) or a non-negative deviation (Panel *b*) from the punisher's contribution. US = United States; UK = United Kingdom; MA = Morocco; TR = Turkey. The error bars show ±1 *SEM* clustered on individuals; *p*-values from Kruskal-Wallis and (pooled) Mann-Whitney tests.

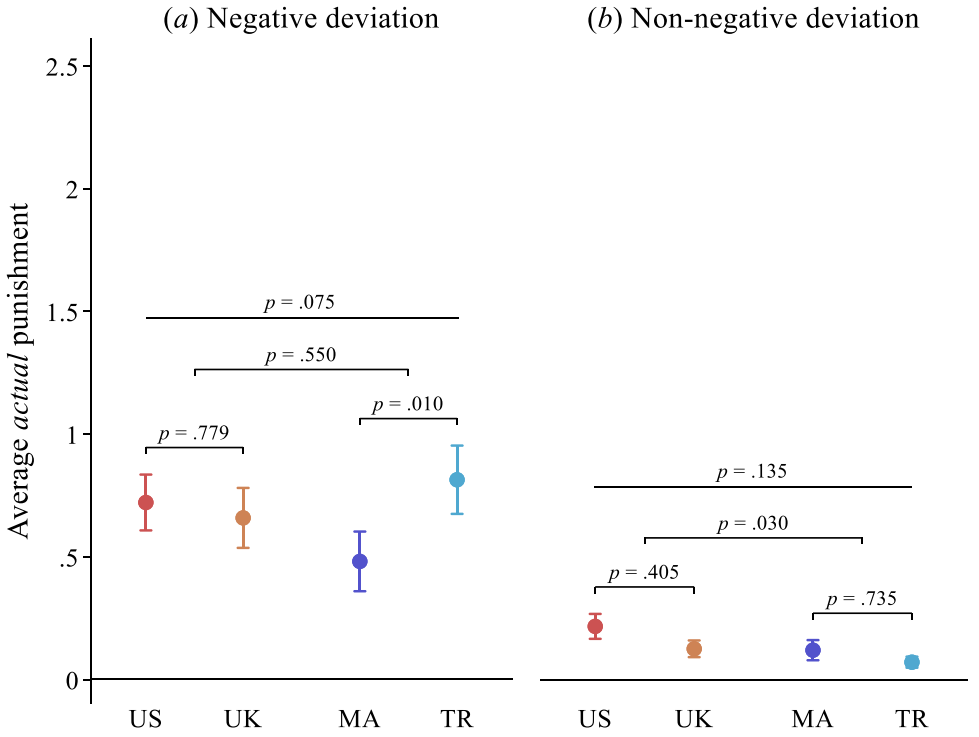

Fig. 4. Average *actual* punishment points by country for a negative deviation (Panel *a*) or a non-negative deviation (Panel *b*) from the punisher's contribution. US = United States; UK = United Kingdom; MA = Morocco; TR = Turkey. The error bars show ±1 *SEM* clustered on individuals; *p*-values from Kruskal-Wallis and (pooled) Mann-Whitney tests.

5.2. Actual punishment

Actual punishment indicates a preference ("willingness to pay") to impose costs on a group member. Fig. 4 shows actual punishment, again distinguishing between negative and non-negative deviations of the target from the punisher's contribution (i.e., prosocial, and antisocial punishment). Actual prosocial punishment differs weakly significantly across the four subject pools. However, we find statistically similar levels of actual prosocial punishment across the UU and MT clusters, and within the UU cluster when comparing the US and UK samples. Within the MT cluster, actual prosocial punishment is significantly higher in the TR sample than in the MA sample.

Antisocial punishment is statistically similar across the four samples. Although the differences appear small, we find significantly higher antisocial punishment in the UU cluster than in the MT cluster. We do not find significant differences within the UU cluster when comparing the US and UK samples, nor within the MT cluster when comparing the MA and TR samples. Tobit regression analysis broadly confirms these results (Online Appendix G, Table G2).

We summarize these findings as:

Result 2a. *Actual punishment of free riding is similar across the two clusters. Antisocial punishment is slightly higher in the UU than the MT cluster.*

These results partially reject Hypothesis 2a. As predicted, we find no differences in actual punishment of free riding. However, contrary to the prediction of Hypothesis 2a, we find slightly lower antisocial punishment in the MT cluster than in the UU cluster.

Our final step is to explain contributions in the P-Game and to test Hypothesis 2b, which predicts higher cooperation in the UU than the MT cluster.

5.3. Explaining contributions in the P-Game

Fig. 5a illustrates average beliefs about others' contributions (b^s) per subject pool s in the P-Game. Beliefs differ significantly between the four subject pools. Beliefs are significantly higher in the UU cluster than in the MT cluster. Within clusters, beliefs are weakly significantly higher in the UK than in the US sample and statistically similar in the MA and TR samples.[21]

Fig. 5b shows the average contributions in the P-Game which differ significantly between the four samples. We also find significantly higher contributions in the UU cluster compared to the MT cluster. A comparison within the UU cluster shows significantly higher contributions in the UK sample compared to the US sample, but we find no differences within the MT cluster.

We also compare contributions in the N-Game with contributions in the P-Game. Interestingly, contributions in the P-Game are significantly higher than in the N-Game for the US (8.06 vs. 7.10, $Z_{US} = -2.42$, $p_{US} = 0.015$) and the UK (10.39 vs. 7.88, $Z_{UK} = -3.65$, $p_{UK} < 0.001$), but not for Morocco (5.36 vs. 4.63, $Z_{MA} = -0.78$, $p_{MA} = 0.438$) and Turkey (5.47 vs. 5.65, $Z_{TR} = -0.45$, $p_{TR} = 0.658$; all tests are Wilcoxon signed-rank tests).

We summarize these observations in the next result, which supports Hypothesis 2b:

Result 2b. *In the presence of punishment, beliefs and contributions are higher in the UU than in the MT cluster.*

Next, we identify and test for two channels through which punishment can affect cooperation: First, participants might expect others to contribute more when punishment is available (i.e., higher cooperative beliefs in the P-Game than in the N-Game), leading to increased contributions from (imperfect) conditional cooperators. Yet, we find relatively weak support for this channel. Beliefs in the P-Game appear statistically similar to the N-Game for the US (8.89 vs. 8.88, $Z_{US} = 0.39$, $p_{US} = 0.702$), Morocco (6.95 vs. 6.22, $Z_{MA} < 0.01$, $p_{MA} = 0.999$) and the Turkish sample (8.22 vs. 8.38, $Z_{TR} = 0.96$, $p_{TR} = 0.342$), and only weakly significantly higher for the UK sample (10.08 vs. 8.72, $Z_{UK} = -1.72$, $p_{UK} = 0.086$; all tests are Wilcoxon signed-rank tests). The share of participants who report a higher belief in the P-Game relative to the N-Game is 34 % in the US, 41 % in the UK, 31 % in Morocco and 31 % in the Turkish sample, with most participants reporting the same or a lower belief. Yet, participants who report a higher belief in the P-Game are also more likely to raise their contribution as changes in beliefs and contributions are positively and significantly correlated in all but the MA sample (Spearman's $\rho_{US} = 0.45$, $p_{US} < 0.001$; $\rho_{UK} = 0.53$, $p_{UK} < 0.001$; $\rho_{MA} = -0.01$, $p_{MA} = 0.940$; $\rho_{TR} = 0.31$, $p_{TR} = 0.004$).

Second, punishment reduces the expected payoff from free riding and thus induces participants to closely match what they expect others to contribute. The importance of beliefs—relative to cooperative attitudes—in explaining actual cooperation thus increases as imperfect conditional cooperators and free riders align their actual contribution more closely with their belief than suggested by their contribution schedules. To test for this channel, we apply the first behavioral mechanism to explain contributions as a function of cooperative attitudes in the N-Game and beliefs in the P-Game $c_i^s = f(a_i^s(b_i^s))$. The resulting predicted contributions (indicated by diamonds (◆) in Fig. 5b) vary significantly across the four samples (Kruskal-Wallis test, $\chi^2(3) = 13.83$, $p = 0.003$). Predicted contributions in the UU cluster are significantly higher than in the MT cluster (pooled Mann-Whitney $Z = 3.68$, $p < 0.001$). There are no differences in our within-cluster analysis of predicted contributions when comparing the UK and US samples (Mann-Whitney $Z = -0.91$, $p = 0.362$) or the MA and TR samples (Mann-Whitney $Z = 0.25$, $p = 0.806$). However, the predicted contributions are significantly lower than the actual contributions for all samples (Wilcoxon signed-rank test $Z_{US} = 4.02$, $p_{US} < 0.001$; $Z_{UK} = 4.89$, $p_{UK} < 0.001$; $Z_{MA} = 1.73$, $p_{MA} = 0.084$; $Z_{TR} = 1.95$, $p_{TR} = 0.051$).

To explain the gap between actual and predicted contributions in the P-Game, we regress, in analogous fashion to Table 3,

[21] Accuracy of beliefs—defined as the deviation from the actual contribution of others ($c_i^s - b_i^s$)—across the four subject pools is higher in the UU cluster than in the MT cluster (average accuracy is 0.65 in the UU cluster vs. 2.13 in the MT cluster). The difference in inaccuracy between clusters is significant, whereas differences within clusters are not (Table D.1 in the Online Appendix).

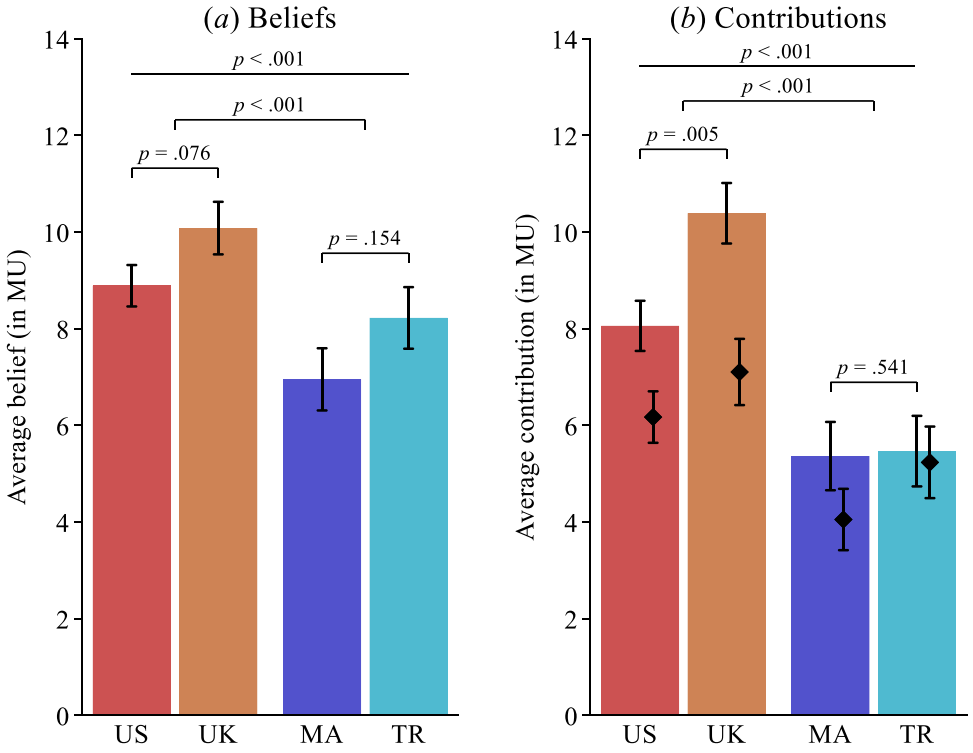

Fig. 5. Average beliefs about others' contributions (Panel *a*) and contributions to the public good (Panel *b*) in the P-Game by country. Diamonds (◆) indicate the average predicted contributions based on cooperative attitudes (from the N-Game) and beliefs. US = United States; UK = United Kingdom; MA = Morocco; TR = Turkey. Error bars indicate ±1 *SEM; p*-values from Kruskal-Wallis and (pooled) Mann-Whitney tests.

contributions (c_i^s) on predicted contributions ($\widehat{c_i^s} = a_i^s(b_i^s)$), beliefs ($b_i^s$), the MT cluster dummy and interaction terms pooling the four samples (Table 4, Col. 1). The coefficients "Predicted contribution" and "Belief" are highly significant, and the coefficient for "Belief" appears larger than the coefficient in Table 3. This confirms the importance of the second channel and implies that, in the presence of punishment, beliefs become more important for explaining contributions. Matching the expected contribution of others may be a strategy to avoid punishment in this game. The weakly significant positive coefficient "Predicted contribution × MT cluster" and the highly significant negative coefficient "Belief × MT cluster" indicate differences in the effect of predicted contributions and beliefs across clusters.

The within-cluster analyses of the UU cluster reveal no significant differences in the effect of predicted contributions, beliefs or level differences in conditional contributions when comparing the US and UK samples (insignificant interactions between "Predicted contributions" or "Belief" and the respective country dummy; insignificant country dummies; Table 4, Col. 2). Similarly, we find no significant within-cluster variation in the effect of predicted contributions, beliefs or level differences within the MT cluster ($F(1, 342) < 0.01$, $p = 0.967$; $F(1, 342) = 0.03$, $p = 0.863$; $F(1, 342) = 1.03$, $p = 0.310$; resp.). A simulation analysis confirms the increased importance of beliefs—relative to cooperative attitudes—in explaining cluster differences in contributions when punishment is available (Online Appendix I). Here, 38 % of the cluster difference can be attributed to beliefs, but only 5 % to cooperative attitudes.

5.4. Discussion

Punishment in a one-shot game is a costly expression of a dislike of others' contributions that cannot be rationalized as a strategic investment because there are no future interactions. Nevertheless, our subjects did punish those group members who contributed less than them. However, surprisingly, antisocial punishment was lower in the MT subject pools than in the UU subject pools, partially rejecting Hypothesis 2a. Yet, these differences in antisocial punishment pools are small. Like in the situation without punishment, the biggest difference we observe is regarding beliefs about others' cooperativeness, where we confirm Hypothesis 2b of higher beliefs and higher cooperation in the UU cluster than the MT cluster.

6. Concluding discussion

This paper contributes to our understanding of behavioral mechanisms that explain voluntary cooperation in one-shot public goods games. We present two behavioral mechanisms and apply them to test their cross-societal generalizability: (1) cooperative attitudes and beliefs about others' cooperativeness jointly explain cooperation (the ABC methodology); and (2) peer punishment. In public

Table 4
Explaining unconditional contributions in the P-Game.

Dependent variable: unconditional contribution	(1) By cluster	(2) By country
Predicted contribution	0.205***	0.277***
	(0.059)	(0.089)
Belief	0.828***	0.732***
	(0.081)	(0.125)
Predicted contribution × MT cluster	0.231*	
	(0.120)	
Belief × MT cluster	−0.477***	
	(0.137)	
MT cluster	0.580	
	(0.858)	
Predicted contribution × UK		−0.147
		(0.122)
Belief × UK		0.164
		(0.176)
UK		0.540
		(1.055)
Predicted contribution × MA		0.160
		(0.166)
Belief × MA		−0.395*
		(0.205)
MA		1.266
		(1.236)
Predicted contribution × TR		0.169
		(0.164)
Belief × TR		−0.357*
		(0.194)
TR		−0.056
		(1.100)
Socio-economic controls	Yes	Yes
Constant	−2.930	−2.488
	(2.223)	(2.493)
R^2	0.59	0.60
No. obs.	360	360

Notes: OLS estimates with robust *SE* in parentheses. UU cluster: UK and US; MT cluster: Morocco and Turkey. Socio-economic control variables: age, female, urban background, middle class, single child, economics/business student. None is significant at any conventional level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

goods experiments conducted in the UK and the US, and two culturally, economically, and institutionally different societies, Morocco and Turkey, we found that conditionally cooperative attitudes (as a proxy for cooperative preferences) are mostly similar across the four subject pools, and it is differences in beliefs that explain a significant fraction of the variation in cooperative behavior. In a second peer punishment experiment, we found strong differences in how participants responded to the presence of punishment. Only participants in the UK and US significantly increased their contributions in the game with punishment compared to the game without punishment. We found that the actual punishment was similar in all four samples.

Our results have several implications for understanding human cooperation. First, conditional cooperation is prevalent in all four subject pools studied here (and previously – see Section 2) and is thus not restricted to US and Western European subject pools, where most of the previous related research has been conducted. Second, our results from the public goods experiments with and without punishment suggest that the observed differences in cooperation across societies are only to a small extent due to differences in the distribution of cooperative preferences across societies but are mostly driven by differences in beliefs about the cooperativeness of others. We thus provide a proof of concept for the importance of beliefs in explaining societal differences, albeit in the relatively narrow context of our public goods game. Our results highlight the importance of beliefs about others' cooperativeness and how they may be a channel through which voluntary cooperation can be promoted in different societies, which also has practical implications. Beliefs are likely more malleable than one's own cooperative preferences, which are likely to be stable over time (e.g., Bruhin et al. (2019); Chuang and Schechter (2015); Carlsson et al. (2014); Volk et al. (2012); Gächter et al. (2022)).

Third, we found very little antisocial punishment (close to zero in Morocco and Turkey, and slightly positive in the UK and US). This is interesting because it contradicts previous research reporting substantial antisocial punishment across societies in repeated games (Herrmann et al., 2008) and one-shot games (Gächter and Herrmann, 2009; Gächter and Herrmann, 2011). In repeated games with peer punishment, several mechanisms can explain the observed pattern of antisocial punishment. First, the punishment of high contributors who do not fully contribute to the public good might result from a strategic intention to maintain or increase the contributions of others. Second, antisocial punishment can result from the desire of a low contributor to retaliate pre-emptively against expected punishment from a higher contributor. Third, antisocial punishment might be an expression of inequality aversion (Thöni, 2014). Finally, antisocial punishment can be revenge for previous game play and thus take on the characteristics of a multi-round feud (Nikiforakis and Engelmann, 2011). The one-shot design of the present study excludes any strategic incentive to punish, does not allow for within-period retaliatory punishment (since there is only one punishment stage), and excludes the possibility of retaliation across

periods. Therefore, we interpret punishment as an elicitation of a preference for sanctioning group members, mostly free riding behavior. We found no systematic variation in such preferences across our culturally diverse subject pools.

It is important to note that for some of our measures we find greater within-cluster variation for the Morocco and Turkey cluster compared to the US and UK cluster. While our analysis does not require that the within-cluster variation is similar, the question arises as to what explains this variation. One candidate is substantial differences in positive and negative reciprocity (as measured by Falk et al. (2018)), which may be particularly important for behavior and beliefs in cooperation games (Online Appendix Fig. A2f).

In summary, given that the four countries from which we drew our samples differ along many cultural and institutional dimensions, we can argue for our two main findings: First, differences in beliefs about others' cooperativeness are more important than differences in cooperative preferences in driving differences in cooperation across the subject pools we study here, with two Western societies, the UK and the US, having higher beliefs than the two non-Western societies, Morocco and Turkey. Second, punishment behavior is similar across subject pools, suggesting that people have similar preferences for enforcing cooperation regardless of the society they live in.

We conclude with three remarks about our contribution and future research. First, a main contribution of this paper is methodological: we show that the ABC approach (and punishment) can explain voluntary cooperation, including differences in the level of cooperation across different subject pools. While our evidence is limited to four subject pools, they span a considerable cultural, economic, and institutional distance. Differences in beliefs turn out to be more important than differences in cooperative preferences and punishment. Based on the evidence presented in this paper, we believe that the ABC approach is transferable across different cultural contexts. Future research should use this approach to further investigate the origins of cooperation differences across societies.

Second, our results suggest that future research should apply our behavioral mechanisms in more countries (and ideally with representative subject pools) to obtain more comprehensive evidence on the relative importance of cooperative preferences, beliefs, and punishment across the globe. Based on our findings, we predict less variation in cooperative preferences than in beliefs across societies. Evidence on perceptions of trust from the World Values Survey can be used as a proxy for cooperative beliefs (see the discussion of beliefs in Section 2 and, for example, Thöni et al. (2012)). Since there is little variation in attitudes towards cooperation even between very different societies, perceptions of trust and trustworthiness will shape beliefs and predict relative cooperation in particular societies.

Third, we measure preferences, beliefs, and behavior in the specific context of a public goods game. Therefore, any application might be limited to situations with similar incentive structures. However, our behavioral mechanisms are not limited to the specific context in which we have applied them here. Future research can therefore use our behavioral mechanisms to analyze the behavioral determinants of voluntary cooperation of any group and should investigate whether our findings hold in other contexts.

Author contributions

Conceptualization: SG and TW; Data curation (running the experiments): TW, JS, BB and FLL in the US, the UK, Turkey and Morocco, respectively. Writing – original draft: SG, TW, and JS. Writing – review & editing: BB, FLL.

Declaration of Competing Interest

The authors declare no financial or other interest.

Data availability

Data, software, and analysis codes are available at https://osf.io/m3u7v/.

Acknowledgments

This work was supported by the European Research Council [grant numbers ERC-AdG 295707 COOPERATION; ERC-AdG 101020453 PRINCIPLES] and the Economic and Social Research Council [grant number ES/K002201/1]. We thank an associate editor, the referees, and Stefania Bortoletti, Lucas Molleman, Christian Thöni, Ori Weisel and participants of the Royal Economic Society Conference 2018, the 2019 Stanford Conference on Behavioral Economics and Development, the MENA network conference at NYU Abu Dhabi 2022 and various workshops for their valuable comments.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jebo.2023.09.006.

References

Alesina, A., Giuliano, P., 2015. Culture and institutions. J. Econ. Lit. 53, 898–944.
Algan, Y., Cahuc, P., 2013. Trust and growth. Ann. Rev. Econ. 5, 521–549.

Anderson, C.M., Putterman, L., 2006. Do non-strategic sanctions obey the law of demand? The demand for punishment in the voluntary contribution mechanism. Game. Econ. Behav. 54, 1–24.
Apicella, C., Norenzayan, A., Henrich, J., 2020. Beyond WEIRD: A review of the last decade and a look ahead to the global laboratory of the future. Evol. Hum. Behav. 41, 319–329.
Arechar, A.A., Gächter, S., Molleman, L., 2018. Conducting interactive experiments online. Exper. Econ. 21, 99–131.
Balliet, D., Van Lange, P.A.M., 2013. Trust, conflict, and cooperation: a meta-analysis. Psychol. Bull. 139, 1090–1112.
Bigoni, M., Bortolotti, S., Casari, M., Gambetta, D., 2019. At the root of the north-south cooperation gap in Italy: preferences or beliefs? Econ. J. 129, 1139–1152.
Bigoni, M., Bortolotti, S., Casari, M., Gambetta, D., Pancotto, F., 2016. Amoral familism, social capital, or trust? The behavioural foundations of the Italian north-south divide. Econ. J. 126, 1318–1341.
Bilancini, E., Boncinelli, L., Celadin, T., 2022. Social value orientation and conditional cooperation in the online one-shot public goods game. J. Econ. Behav. Organ. 200, 243–272.
Bisin, A., Verdier, T., et al., 2011. Chapter 9 - the economics of cultural transmission and socialization. In: Benhabib, J., et al. (Eds.), Handbook of Social Economics. North-Holland, pp. 339–416.
Blanco, M., Engelmann, D., Koch, A., Normann, H.-T., 2010. Belief elicitation in experiments: is there a hedging problem? Exper. Econ. 13, 412–438.
Bohnet, I., Greig, F., Herrmann, B., Zeckhauser, R., 2008. Betrayal aversion. Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States. Am. Econ. Rev. 98, 294–310.
Bohnet, I., Herrmann, B., Zeckhauser, R., 2010. Trust and the reference points for trustworthiness in Gulf and Western countries. Q. J. Econ. 125, 811–828.
Boosey, L., Mark Isaac, R., Norton, D., Stinn, J., 2020. Cooperation, contributor types, and control questions. J. Behav. Exper. Econ. 85, 101489.
Boyd, R., Richerson, P.J., 1985. Culture and the Evolutionary Process. University of Chicago Press, Chicago.
Boyd, R., Richerson, P.J., 2009. Culture and the evolution of human cooperation. Philosoph. Transact. Roy. Soc. B: Biolog. Sci. 364, 3281–3288.
Brislin, R.W., 1970. Back-translation for cross-cultural research. J. Cross-Cult. Psychol. 1, 185–216.
Brosig-Koch, J., Helbach, C., Ockenfels, A., Weimann, J., 2011. Still different after all these years: Solidarity behavior in East and West Germany. J. Public Econ. 95, 1373–1376.
Bruhin, A., Fehr, E., Schunk, D., 2019. The many faces of human sociality: uncovering the distribution and stability of social preferences. J. Eur. Econ. Assoc. 17, 1025–1069.
Bruhin, A., Janizzi, K., Thöni, C., 2020. Uncovering the heterogeneity behind cross-cultural variation in antisocial punishment. J. Econ. Behav. Organ. 180, 291–308.
Carlsson, F., Johansson-Stenman, O., Nam, P.K., 2014. Social preferences are stable over long periods of time. J. Public Econ. 117, 104–114.
Carpenter, J.P., 2007. The demand for punishment. J. Econ. Behav. Organ. 62, 522–542.
Casari, M., 2005. On the design of peer punishment experiments. Exper. Econ. 8, 107–115.
Chaudhuri, A., 2011. Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. Exper. Econ. 14, 47–83.
Choi, S., Kim, B.-Y., Lee, J., Lee, S., 2020. A tale of two Koreas: Property rights and fairness. J. Econ. Behav. Organ. 170, 112–130.
Chuah, S.-H., Gächter, S., Hoffmann, R., Tan, J.H.W., 2023. Who discriminates? Evidence from a trust game experiment across three societies. J. Econ. Psychol. 97, 102630.
Chuah, S.H., Gächter, S., Hoffmann, R., Tan, J.H.W., 2016. Religion, discrimination and trust across three cultures. Eur. Econ. Rev. 90, 280–301.
Chuang, Y., Schechter, L., 2015. Stability of experimental and survey measures of risk, time, and social preferences: a review and some new results. J. Develop. Econ. 117, 151–170.
Cohn, A., Maréchal, M.A., Tannenbaum, D., Zünd, C.L., 2019. Civic honesty around the globe. Science 365, 70–73.
Croson, R., 2007. Theories of commitment, altruism and reciprocity: evidence from linear public goods games. Econ. Inq. 45, 199–216.
Cubitt, R., Drouvelis, M., Gächter, S., 2011. Framing and free riding: emotional responses and punishment in social dilemma games. Exper. Econ. 14, 254–272.
Egas, M., Riedl, A., 2008. The economics of altruistic punishment and the maintenance of cooperation. Proceed. Roy. Soc. B - Biolog. Sci. 275, 871–878.
Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., Sunde, U., 2018. Global evidence on economic preferences. Q. J. Econ. 133, 1645–1692.
Fehr, E., 2009. The economics and biology of trust. J. Eur. Econ. Assoc. 7, 235–266.
Fehr, E., Gächter, S., 2002. Altruistic punishment in humans. Nature 415, 137–140.
Fehr, E., Schurtenberger, I., 2018. Normative foundations of human cooperation. Nat. Hum. Behav. 2, 458–468.
Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Exper. Econ. 10, 171–178.
Fischbacher, U., Gächter, S., 2010. Social preferences, beliefs, and the dynamics of free riding in public good experiments. Am. Econ. Rev. 100, 541–556.
Fischbacher, U., Gächter, S., Fehr, E., 2001. Are people conditionally cooperative? Evidence from a public goods experiment. Econ. Lett. 71, 397–404.
Fischbacher, U., Gächter, S., Quercia, S., 2012. The behavioral validity of the strategy method in public good experiments. J. Econ. Psychol. 33, 897–913.
Fisman, R., Miguel, E., 2007. Corruption, norms, and legal enforcement: evidence from diplomatic parking tickets. J. Politi. Econ. 115, 1020–1048.
Fosgaard, T.R., Hansen, L.G., Wengström, E., 2014. Understanding the nature of cooperation variability. J. Public Econ. 120, 134–143.
Frey, B.S., Meier, S., 2004. Social comparisons and pro-social behavior. Testing 'conditional cooperation' in a field experiment. Am. Econ. Rev. 94, 1717–1722.
Gächter, S., 2007. Conditional cooperation: Behavioral regularities from the lab and the field and their policy implications. In: Frey, B.S., Stutzer, A. (Eds.), Psychology and Economics: A Promising New Cross-Disciplinary Field (CESifo Seminar Series. The MIT Press, Cambridge, pp. 19–50.
Gächter, S., 2010. (Dis)advantages of student subjects: What is your research question? Behav. Brain Sci. 33, 92–93.
Gächter, S., Herrmann, B., 2009. Reciprocity, culture, and human cooperation: previous insights and a new cross-cultural experiment. Philosoph. Transact. Roy. Soc. B – Biolog. Sci. 364, 791–806.
Gächter, S., Herrmann, B., 2011. The limits of self-governance when cooperators get punished: experimental evidence from urban and rural Russia. Eur. Econ. Rev. 55, 193–210.
Gächter, S., Herrmann, B., Thöni, C., 2004. Trust, voluntary cooperation, and socio-economic background: survey and experimental evidence. J. Econ. Behav. Organ. 55, 505–531.
Gächter, S., Herrmann, B., Thöni, C., 2010. Culture and cooperation. Philosoph. Transact. Roy. Soc. B: Biolog. Sci. 365, 2651–2661.
Gächter, S., Kölle, F., Quercia, S., 2017. Reciprocity and the tragedies of maintaining and providing the commons. Nat. Hum. Behav. 1, 650–656.
Gächter, S., Kölle, F., Quercia, S., 2022. Preferences and perceptions in Provision and Maintenance public goods. Games Econ. Behav. 135, 338–355.
Gächter, S., Renner, E., 2010. The effects of (incentivized) belief elicitation in public goods experiments. Exper. Econ. 13, 364–377.
Gächter, S., Renner, E., 2018. Leaders as role models and 'belief managers' in social dilemmas. J. Econ. Behav. Organ. 154, 321–334.
Gächter, S., Schulz, J.F., 2016. Intrinsic honesty and the prevalence of rule violations across societies. Nature 531, 496–499.
Greif, A., 1994. Cultural beliefs and the organization of society: A historical reflection on collectivist and individualist societies. J. Polit. Econ. 102, 912–950.
Greiner, B., 2015. Subject pool recruitment procedures: organizing experiments with ORSEE. J. Econ. Sci. Assoc. 1, 114–125.
Guiso, L., Sapienza, P., Zingales, L., 2006. Does culture affect economic outcomes? J. Econ. Perspect. 20, 23–48.
Guiso, L., Sapienza, P., Zingales, L., 2008. Social capital as good culture. J. Eur. Econ. Assoc. 6, 295–320.
Hauk, E., Saez-Marti, M., 2002. On the cultural transmission of corruption. J. Econom. Theory 107, 311–335.
Henrich, J., 2015. The Secret of Our Success: How Culture is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter. Princeton University Press, Princeton.
Henrich, J., 2020. The WEIRDest People in the World: How the West Became Psychologically Peculiar and Particularly Prosperous. Farrar, Strauss and Giroux, New York.
Henrich, J., Boyd, R., Bowles, S., Camerer, C.F., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Henrich, N., Hill, K., Gil-White, F., Gurven, M., Marlowe, F.W., Patton, J.Q., Tracer, D., 2005. Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. Behav. Brain Sci. 28, 795–855.
Henrich, J., Heine, S.J., Norenzayan, A., 2010. The weirdest people in the world? Behav. Brain Sci. 33, 61–83.

Herrmann, B., Thöni, C., 2009. Measuring conditional cooperation: a replication study in Russia. Exper. Econ. 12, 87–92.
Herrmann, B., Thöni, C., Gächter, S., 2008. Antisocial punishment across societies. Science 319, 1362–1367.
Hofstede, G.H., Hofstede, G., 2001. Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations. Sage Publications, Thousand Oaks, CA.
Inglehart, R., Haerpfer, C., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., Lagos, M., Norris, P., Ponarin, E., 2014. World Values Survey: Round Five - Country-Pooled Datafile. JD Systems Institute, Madrid. Version: www.worldvaluessurvey.org/WVSDocumentationWV5.jsp (Ed.)ˆ(Eds.).
Inglehart, R., Welzel, C., 2005. Modernization, Cultural Change, and Democracy. The Human Development Sequence. Cambridge University Press, Cambridge.
Isler, O., Gächter, S., Maule, A.J., Starmer, C., 2021. Contextualised strong reciprocity explains selfless cooperation despite selfish intuitions and weak social heuristics. Sci. Rep. 11, 13868.
Kaufmann, D., Kraay, A., 2016. Worldwide Governance Indicators (Ed.)ˆ(Eds.)Dataset retrieved from. http://www.govindicators.org.
Keser, C., van Winden, F., 2000. Conditional cooperation and voluntary contributions to public goods. Scand. J. Econ. 102, 23–39.
Kim, B.-Y., Choi, S., Lee, J., Lee, S., Choi, K., 2017. Do institutions affect social preferences? Evidence from divided Korea. J. Comparat. Econ. 45, 865–888.
Knack, S., Keefer, P., 1997. Does social capital have an economic payoff? A cross-country investigation. Q. J. Econ. 112, 1251–1288.
Kocher, M.G., Cherry, T., Kroll, S., Netzer, R.J., Sutter, M., 2008. Conditional cooperation on three continents. Econ. Lett. 101, 175–178.
Kocher, M.G., Martinsson, P., Matzat, D., Wollbrant, C., 2015. The role of beliefs, trust, and risk in contributions to a public good. J. Econ. Psychol. 51, 236–244.
List, J.A., 2004. Young, selfish and male: field evidence of social preferences. Econ. J. 114, 121–149.
Martinsson, P., Pham-Khanh, N., Villegas-Palacio, C., 2013. Conditional cooperation and disclosure in developing countries. J. Econ. Psychol. 34, 148–155.
Molleman, L., Gächter, S., 2018. Societal background influences social learning in cooperative decision making. Evol. Hum. Behav. 39, 547–555.
Molleman, L., Kölle, F., Starmer, C., Gächter, S., 2019. People prefer coordinated punishment in cooperative interactions. Nat. Hum. Behav. 3, 1145–1153.
Muthukrishna, M., Bell, A.V., Henrich, J., Curtin, C.M., Gedranovich, A., McInerney, J., Thue, B., 2020. Beyond western, educated, industrial, rich, and democratic (WEIRD) psychology: measuring and mapping scales of cultural and psychological distance. Psychol. Sci. 31, 678–701.
Nikiforakis, N., 2008. Punishment and counter-punishment in public good games: can we really govern ourselves? J. Publ. Econ. 92, 91–112.
Nikiforakis, N., Engelmann, D., 2011. Altruistic punishment and the threat of feuds. J. Econ. Behav. Organ. 78, 319–332.
Nikiforakis, N., Normann, H., 2008. A comparative statics analysis of punishment in public goods experiments. Exper. Econ. 11, 358–369.
Oosterbeek, H., Sloof, R., van de Kuilen, G., 2004. Cultural differences in ultimatum game experiments: Evidence from a meta-analysis. Exper. Econ. 7, 171–188.
Prediger, S., Vollan, B., Herrmann, B., 2014. Resource scarcity and antisocial behavior. J. Publ. Econ. 119, 1–9.
Raihani, N.J., Bshary, R., 2019. Punishment: one tool, many uses. Evolut. Hum. Sci. 1, e12.
Romano, A., Balliet, D., Yamagishi, T., Liu, J.H., 2017. Parochial trust and cooperation across 17 societies. Proc. Natl Acad. Sci. 114, 12702–12707.
Roth, A.E., Prasnikar, V., Okuno-Fujiwara, M., Zamir, S., 1991. Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: an experimental study. Am. Econ. Rev. 81, 1068–1095.
Rustagi, D., 2022. Historical self-governance and norms of cooperation. CeDEx Discussion Paper 2022-04, University of Nottingham.
Rustagi, D., Engel, S., Kosfeld, M., 2010. Conditional cooperation and costly monitoring explain success in forest commons management. Science 330, 961–965.
Schulz, J.F., 2022. Kin networks and institutional development. Econ. J. 132, 2578–2613.
Schulz, J.F., Bahrami-Rad, D., Beauchamp, J.P., Henrich, J., 2019. The Church, intensive kinship, and global psychological variation. Science 366, eaau5141.
Sylwester, K., Herrmann, B., Bryson, J.J., 2013. *Homo homini lupus*? Explaining antisocial punishment. J. Neurosci. Psychol. Econ. 6, 167–188.
Tabellini, G., 2008. The scope of cooperation: values and incentives. Q. J. Econ. 123, 905–950.
Thöni, C., 2014. Inequality aversion and antisocial punishment. Theor. Decisi. 76, 529–545.
Thöni, C., 2019. Cross-cultural behavioral experiments: potential and challenges. In: Schram, A., Ule, A. (Eds.), Handbook of Research Methods and Applications in Experimental Economics. Edward Elgar Publishing, Cheltenham.
Thöni, C., Tyran, J.-R., Wengström, E., 2012. Microfoundations of social capital. J. Publ. Econ. 96, 635–643.
Thöni, C., Volk, S., 2018. Conditional Cooperation: Review and Refinement. Econ. Lett. 171, 37–40.
Volk, S., Thöni, C., Ruigrok, W., 2012. Temporal stability and psychological foundations of cooperation preferences. J. Econ. Behav. Organ. 81, 664–676.
Walker, J.M., Halloran, M.A., 2004. Rewards and sanctions and the provision of public goods in one-shot settings. Exper. Econ. 7, 235–247.
Weber, T.O., Weisel, O., Gächter, S., 2018. Dispositional free riders do not free ride on punishment. Nat. Commun. 9, 2390.