

Culture and cooperation

Simon Gächter, Benedikt Herrmann and Christian Thöni

Phil. Trans. R. Soc. B 2010 **365**, 2651-2661

doi: 10.1098/rstb.2010.0135

References

[This article cites 73 articles, 11 of which can be accessed free](#)

<http://rstb.royalsocietypublishing.org/content/365/1553/2651.full.html#ref-list-1>

Rapid response

[Respond to this article](#)

<http://rstb.royalsocietypublishing.org/letters/submit/royptb;365/1553/2651>

Subject collections

Articles on similar topics can be found in the following collections

[behaviour](#) (1389 articles)

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)

To subscribe to *Phil. Trans. R. Soc. B* go to: <http://rstb.royalsocietypublishing.org/subscriptions>

Culture and cooperation

Simon Gächter^{1,2,*}, Benedikt Herrmann¹ and Christian Thöni³

¹*Centre of Decision Research and Experimental Economics, School of Economics, University of Nottingham, Sir Clive Granger Building, University Park, Nottingham NG7 2RD, UK*

²*Center for Economic Studies and Ifo Institute for Economic Research, Institute for the Study of Labor, Munich DE-81679, Germany*

³*University of St Gallen, SEW-HSG, Varnbuelstrasse 14, CH-9000 St Gallen, Switzerland*

Does the cultural background influence the success with which genetically unrelated individuals cooperate in social dilemma situations? In this paper, we provide an answer by analysing the data of Herrmann *et al.* (2008a), who studied cooperation and punishment in 16 subject pools from six different world cultures (as classified by Inglehart & Baker (2000)). We use analysis of variance to disentangle the importance of cultural background relative to individual heterogeneity and group-level differences in cooperation. We find that culture has a substantial influence on the extent of cooperation, in addition to individual heterogeneity and group-level differences identified by previous research. The significance of this result is that cultural background has a substantial influence on cooperation in otherwise identical environments. This is particularly true in the presence of punishment opportunities.

Keywords: human cooperation; punishment; culture; experimental public good games

1. INTRODUCTION

Many important social problems of mankind—from interactions in the workplace to tackling climate change—involve the cooperation of genetically unrelated individuals in situations in which collective welfare is jeopardized by individual self-interest. According to one model of human social behaviour, self-interest is a dominant behavioural force and therefore welfare-enhancing cooperation is doomed to fail, unless well-defined small groups interact indefinitely (which allows for targeted punishment by withdrawing cooperation, see Axelrod (1984), Fudenberg & Maskin (1986), Sigmund (2010)). Numerous behavioural experiments, as for example surveyed in Fehr *et al.* (2002b), Fehr & Fischbacher (2004) and Gächter & Herrmann (2009), and other empirical studies (e.g. Gintis *et al.* 2005) have shown that this prediction is way too pessimistic and that much more cooperation exists than is easily compatible with the selfishness assumption. This is particularly true in the presence of punishment opportunities. Many people are willing to exert costly punishment of people whose behaviour they dislike, even when there is no material benefit whatsoever for doing so. However, recent research also suggests that there is substantial individual heterogeneity with regard to prosocial behaviour, in particular in the context of cooperation (e.g. Fischbacher *et al.* 2001; Kurzban & Houser 2005; Bardsley & Moffatt 2007; Kocher *et al.* 2008; Muller *et al.* 2008; Herrmann & Thöni 2009; Fischbacher & Gächter 2010; see Bergmuller *et al.* 2010, for a discussion of personality and

cooperation). Such inter-individual differences have a potential of explaining aggregate behaviour and group-level differences (e.g. Gächter & Thöni 2005; Kurzban & Houser 2005; Gunnthorsdottir *et al.* 2007; Fischbacher & Gächter 2010) and may play a major role for the stability of cooperation (McNamara & Leimar 2010).

An interesting next approach is to jump from the ‘micro-level’ to the ‘macro-level’ and to ask whether there are also differences in cooperation behaviour across cultural backgrounds. When we speak of the ‘cultural background’ we have in mind those sets of beliefs and values that the majority of people in these societies hold and that get ‘transmitted fairly unchanged from generation to generation’ (Guiso *et al.* 2006, p. 23). In particular, influential social scientists such as Inglehart (1997) and Inglehart & Baker (2000) argue on the basis of data from the World Values Survey that there are distinct cultural areas in the world, reflected in people’s value systems. The question we ask in this paper is whether there are differences in experimentally observed cooperation behaviour across distinct world cultures.¹

To answer this question, we will analyse a dataset of highly comparable cross-cultural experiments conducted by Herrmann *et al.* (2008a) with more than 1100 participants in 16 subject pools from six distinct cultural areas around the world. All participants played finitely repeated public good experiments with and without punishment in stable groups, in a design inspired by Fehr & Gächter (2000). This dataset, which we describe in §2 in more detail, along with our methodology of classifying subject pools according to cultural areas, allows us to disentangle the relative importance of individual heterogeneity, group-level differences and cultural heterogeneity for cooperation. To our knowledge, such an analysis has not been done before.

* Author for correspondence (simon.gaechter@nottingham.ac.uk).

One contribution of 14 to a Theme Issue ‘Cooperation and deception: from evolution to mechanisms’.

In principle, survey methods could also be applied to uncover cross-cultural differences. However, subjects do not have an incentive to admit their true social preferences when it costs nothing to pass for being cooperative and prosocial. When surveyed, presumably only a few people would admit to being selfish. By contrast, behavioural experiments have the advantage that actual behaviour rather than stated intentions is observed. In experiments participants can, depending on their decisions, earn considerable amounts of money. Thus, the laboratory allows observation of real decision-making under controlled circumstances. Moreover, our goal of disentangling individual heterogeneity, group-level differences and cultural variation demands a laboratory experimental approach.²

Why might the cultural background matter at all for cooperation? This is an interesting question because the *Homo economicus* model mentioned above suggests that cultural background does not matter: selfishness is universal. The fact that not all people are selfish has recently inspired theoretical models of social preferences, which take this heterogeneity into account (e.g. Fehr & Schmidt (2006) for a survey). Yet, these models are also mute with respect to the influence of the cultural background. In general, economists, with some exceptions (e.g. Roth *et al.* (1991) in a seminal study) have not been interested in cultural differences. This is now changing (for a succinct survey see Fernández 2008). The reasons are theoretical developments (e.g. Greif 1994; Bowles 1998; Bednar & Page 2007; Guiso *et al.* 2008; Tabellini 2008*b*), and better data, both experimental (e.g. Henrich *et al.* 2001; Oosterbeek *et al.* 2004) and non-experimental (Guiso *et al.* 2006; Fernández 2007; Tabellini 2008*a*). By contrast, psychologists have established many profound differences in human behaviour and thinking across cultures (e.g. the reviews by Markus & Kitayama (1991), Nisbett & Cohen (1996), Cohen (2001), Nisbett (2003), Henrich *et al.* (in press), Heine & Buchtel (2009) and Heine & Ruby (2010)). For example, in a recent paper Henrich *et al.* (in press) show that Western subjects, who are most frequently used in behavioural experiments, are actually often the outlier in the range of observed behaviours. Thus, it is an obvious question whether there are also differences in cooperation behaviour across different world cultures. Moreover, evolutionary psychological approaches predict the possibility of cultural differences because people have an evolved psychology that allows them to attune their behaviour to the norms, expectations and (sanctioning) behaviours of others around them (e.g. Boyd & Richerson 2005; Henrich 2004; Henrich & Henrich 2007; Herrmann *et al.* 2007; Nettle 2009; Tomasello *et al.* 2005; Tomasello 2009; Rendell *et al.* 2010; Gintis in press).

From what we know from numerous experiments, we can speculate about potential behavioural channels of cultural influences. First, in the context of cooperation many experiments have shown that people are conditional cooperators who cooperate more the more they believe others will cooperate (e.g. Croson 2007; Gächter 2007; Fischbacher & Gächter 2010).

Any factor that influences beliefs might also influence cooperation. This is also true of framing effects (e.g. Dufwenberg *et al.* 2006) or, more generally, contextual cues, of which the cultural background is an important example. For example, subjects in a public good experiment in Kenya termed the neutrally framed experiment as ‘harambee’, their word for community work (Henrich *et al.* 2005). The way naturally occurring cooperation problems are normally solved in society might influence people’s beliefs about how others will behave. Second, from experiments in which punishment was possible, we know that substantial differences in punishment across subject pools in different cultures can exist and even be anticipated prior to any experience in the particular situation (e.g. Gächter *et al.* (2005) and Gächter & Herrmann (2009) who ran experiments in Russia and Switzerland). Consistent with this observation, Herrmann *et al.* (2008*a*), in experiments which we shall analyse in detail below, showed a large diversity of punishment patterns across different subject pools around the world, resulting in vastly different cooperation levels.

We are of course not the first to investigate cultural influences on cooperation behaviour or prosociality in general (e.g. Oosterbeek *et al.* 2004). Particularly noteworthy are the seminal large-scale studies conducted in small-scale societies around the world (Henrich *et al.* 2010).³ While Henrich and his co-workers (Henrich *et al.* 2001, 2005, 2006) mostly used simple bargaining games and conducted their experiments with members of small-scale societies, the experiments we shall analyse were all conducted in large-scale developed societies. The small-scale societies differ among each other in the extent to which cooperation is important for economic production (e.g. cooperative whale-hunting versus individual hunting and gathering); how strong market integration is (how many calories are bought on the market?); the size of communities; and adherence to a world religion (Henrich *et al.* 2010). Differences on these dimensions explain a large part of the variation that is observed in experimental bargaining games in these small-scale societies (Henrich *et al.* 2005, 2010). Modern developed societies hardly differ on the dimensions of market integration and reliance on cooperation, for all modern societies know division of labour and trade between non-kin (Richerson & Boyd 1999). Thus, in comparison with the small-scale societies, the cultural influence we identify does not come from fundamental differences in socio-economic structures but from historical, religious, political and value differences, which Inglehart & Baker’s (2000) classification of cultural areas around the world, or Hofstede’s (2001) ‘cultural dimensions’ try to capture.

Another distinguishing feature of our approach from previous cross-cultural economics experiments is that many of them test specific (proximate) hypotheses that are derived from the compared cultures (Yamagishi 1988; Yamagishi & Yamagishi 1994; Kachelmeier & Shehata 1997; Yamagishi *et al.* 1998; Hayashi *et al.* 1999; Buchan *et al.* 2002, 2009; Holm & Danielson 2005; Chuah *et al.* 2007, 2009; Bohnet *et al.* 2008, 2010; Wu *et al.* 2009; Bornhorst *et al.* in press).

Our approach is different since our goal is to understand a more fundamental issue—do we find evidence that comparable subjects from modern developed societies that are characterized by large-scale cooperation but differ strongly with regard to historical and cultural values behave differently in games of cooperation? This question is motivated by evolutionary theories of cooperation (Sober & Wilson 1998; Henrich 2004; Nowak 2006; Henrich & Henrich 2007) rather than proximate mechanisms of cultural differences.

The typical methodology of cross-cultural experiments is to observe a comparable subject pool in different societies. The idea is to run experiments in a way that minimizes variations owing to subject pool composition or experimental procedures. In this way any differences that might be observed between cross-societal subject pools are probably due to differences in the cultural background of the compared societies. Our methodology, which we describe in more detail in the next section, builds on this idea but refines it in two ways. First, the data of Herrmann *et al.* (2008a) were collected in six distinct cultural areas according to Inglehart & Baker (2000) and Hofstede (2001). Thus, rather than comparing two cultures, we compare six cultures. Second, we do not identify culture by nationality, because different nations can share largely similar cultural backgrounds. The cultural classification of Inglehart & Baker (2000) gives us at least two different societies in each of the six cultures; in three cultural areas we have data from subject pools from three different societies and in three cultural areas from two different societies. In one culture, 'Protestant Europe', we have data from four subject pools from three countries (in Switzerland we have data from two subject pools, St Gallen and Zurich). This structure of our data allows us to compare within-cultural variation with between-cultural variation, which is impossible if there is only one subject pool per society or cultural area.⁴

Our main findings are that cooperation within cultures is largely similar while there exist highly significant differences between cultures. This is true in public good experiments with and without punishment and also holds for punishment behaviour. This dual observation of within-culture similarity and cross-cultural heterogeneity is the main support for the claim that there are cultural influences on cooperation.

2. THE DATA AND OUR APPROACH

In the following, we first describe the most important details of the design of Herrmann *et al.* (2008a), followed by the details of our classification of cultural areas. Our third step is a description of our main statistical approach for discerning the importance of the cultural background for cooperation and punishment.

We start with the details of the experimental design, which was motivated by the observations from Ostrom *et al.* (1992), Fehr & Gächter (2000) and Fehr & Gächter (2002) who showed that the punishment mechanism has dramatic impacts on contributions in the public goods game. All subjects took part in two experiments, each lasting for 10 periods. The first

experiment always was a public good experiment with no punishment opportunities (we call this the 'N-condition'). The second experiment was a public good experiment with a punishment opportunity (the 'P-condition'). Both experiments were played in stable groups of four subjects for 10 periods. In both experiments, subjects received an endowment of 20 ECU (experimental currency unit) in each period. All subjects decided simultaneously how many ECU they wanted to contribute to a public good. All contributions in a group were summed up and multiplied by 1.6. The resulting amount of ECU was divided *equally* among all subjects in the group. A subject's payoff consisted of the ECUs he or she did not contribute plus his share of the public good. In the N-condition, the stage game ended here and subjects moved on to the next period. Note that in this game it is individually rational (assuming selfish preferences) to contribute nothing to the public good: for every unit contributed a subject earns only 0.4 units in return. However, joint income is maximized if all subjects contribute their entire endowment to the public good. This is due to the fact that the *social return* of contributing is 1.6 per unit contributed.

In the P-condition, there was an additional stage where subjects could reduce each others' incomes at their own cost. All subjects learned the contributions of all other group members. Subjects could then assign punishment points to each other group member. Each punishment point reduced the income of the punished group member by three ECUs. However, punishment was also costly to the punisher. Each punishment point cost the punisher one ECU. For further details, the procedures and the instructions, we refer the reader to Herrmann *et al.* (2008b).

Herrmann *et al.* (2008a) ran these experiments in 16 different locations with a total of 1120 participants. The locations are all over the developed world and span a large set of cross-societal differences (see Herrmann *et al.* (2008b) for the details). As explained above, subjects interacted in stable groups of four members throughout the entire experiment. Therefore, groups constitute the independent units of observations on which all our non-parametric tests will be based. In total, we have data from 280 groups.

Herrmann *et al.* (2008a) designed and ran their experiments in a way that minimizes differences in behaviour that come from subject pool composition or experimental procedures. To ensure this, participants were all undergraduates and thereby very similar with regard to age, education and their socio-economic situation in their respective society. Gender composition was also similar in most subject pools. Thus, any variation we observe between subject pools or cultural regions are unlikely owing to differences in subject pool composition. Similarly, to minimize behavioural variability as introduced by experimental procedures, Herrmann *et al.* (2008a) followed standard practices of cross-cultural experiments as introduced to experimental economics by Roth *et al.* (1991). A detailed discussion of these issues can be found in Herrmann *et al.* (2008b).

An important conceptual step for our purposes is to classify locations into cultural regions according to

Table 1. Cultural classification and number of observations of the cities where our data stems from. Classification taken and adapted from Inglehart & Baker (2000) and Hofstede (2001) (for Southern Europe and Arabic speaking).

city	country	culture	no. of subjects
Boston	USA	English speaking	56
Nottingham	UK		56
Melbourne	Australia		40
Copenhagen	Denmark	Protestant Europe	68
Bonn	Germany		60
Zurich	Switzerland		92
St Gallen	Switzerland		96
Minsk	Belarus	Orthodox/	68
Dnipropetrovs'k	Ukraine	ex-Communist	44
Samara	Russia		152
Athens	Greece	southern Europe	44
Istanbul	Turkey		64
Riyadh	Saudi Arabia	Arabic speaking	48
Muscat	Oman		52
Seoul	South Korea	Confucian	84
Chengdu	China		96

cultural proximity. To avoid being arbitrary, we rely on seminal research by Inglehart (1997) and Inglehart & Baker (2000), who used data from the World Values Survey to identify clusters in world cultures. According to Inglehart & Baker (2000), societies can be characterized by two dimensions: 'traditional versus secular-rational values' and 'survival versus self-expression values'. The first refers to people's attitudes on topics like abortion, national pride, obedience and respect for authorities; the latter refers to attitudes on the importance of economic and physical security over self-expression and quality-of-life; homosexuality, happiness and trust. Table 1 shows the countries where our data stems from and their cultural classification. Where available, we take the classification from the Global Cultural Map (Inglehart & Baker 2000, p. 29 fig. 1). This allows us to classify all countries in the cultural areas 'English speaking', 'Protestant Europe', 'Orthodox/ex-Communist' and 'Confucian'. Among the four remaining countries, only Turkey appears in Inglehart & Baker (2000). An alternative source of information about cultural differences are the four cultural dimensions (power distance, individualism, masculinity and uncertainty avoidance) defined by Hofstede (2001). Using these four dimensions strongly suggests pairing Greece and Turkey. If we calculate the Euclidian distance, then Turkey is the third closest country to Greece in a sample of 71 countries (and the closest one in our sample of countries). Finally, we group the two Arabic subject pools into the category 'Arabic-speaking'.

Before we continue, a caveat is in order. Classifications are always to some extent open to criticism, and Inglehart & Baker (2000) are aware of this (see their discussion on pp. 32–40). We believe, however, that this classification makes a lot of sense, in particular because the identified cultural clusters all share some common history and four of the clusters also

share a common language. Moreover, the identified clusters are also similar with regard to other measures of cultural similarity, such as Hofstede's four cultural dimensions (Hofstede 2001), or norms of civic cooperation, the strength of the rule of law or democracy (see Herrmann *et al.* 2008b, in particular table S1). There is no detailed information on the Arabic countries, but Hofstede groups them under 'Arab world' (Hofstede's sample does not include Oman, but its neighbouring states Saudi Arabia and UAE).

Our main interest is in whether there are cultural differences in contribution decisions and how important they are, if they exist. To analyse these questions our empirical strategy will be twofold. We first describe the data using graphical tools and non-parametric tests to analyse whether there are cultural differences, that is, systematic patterns of different contributions to the public good according to the cultural areas defined above. Cultural differences exist whether the variation between cultures is larger than the variation within cultures. Therefore, we will provide tests of behaviour within a culture as well as tests between cultures. If behaviour is very homogeneous within the culture but different across cultures, we should not find statistically significant differences within the culture but significant differences between cultures. Notice, however, that homogeneity within the culture and differences across cultures are only sufficient for the existence of cultural differences. Significant between-cultural differences can still exist even if there are significant within-cultural differences, provided the within-cultural differences are 'small enough' relative to the between-cultural differences.

The existence of cultural differences does not yet tell us how 'big' they are, also relative to the importance of individual variation and variation that is due to differences between groups. For that purpose we use a nested analysis of variance (ANOVA) model to attribute the amount of variance in the contributions explained by cultural variations, group differences and individual heterogeneity. Our basic linear model underlying the ANOVA uses the exogenous variables *Period*, *Culture*, *Group* and *Individual* (*Period* is the period number, *Culture* is a categorical variable to identify the six cultural clusters, and *Individual* (*Group*) is a dummy variable for each individual (group)). *Individual* is nested in *Group* and *Group* is nested in *Culture*. We use the ANOVA to disentangle the coefficient of determination to separate the explanatory power of our exogenous variables in the N- and the P-condition.⁵ Our approach not only allows us to measure the explanatory power of cultural variation, but also allows us to compare the importance of cultural variation relative with individual and group influences.

3. RESULTS

The main results of the first part of our analysis, which concerns the existence of culture effects, are contained in figures 1–4. Recall that we argued that cultural differences in contributions exist if contributions are more similar within a culture than between cultures. In our analysis, we separate the data according to the

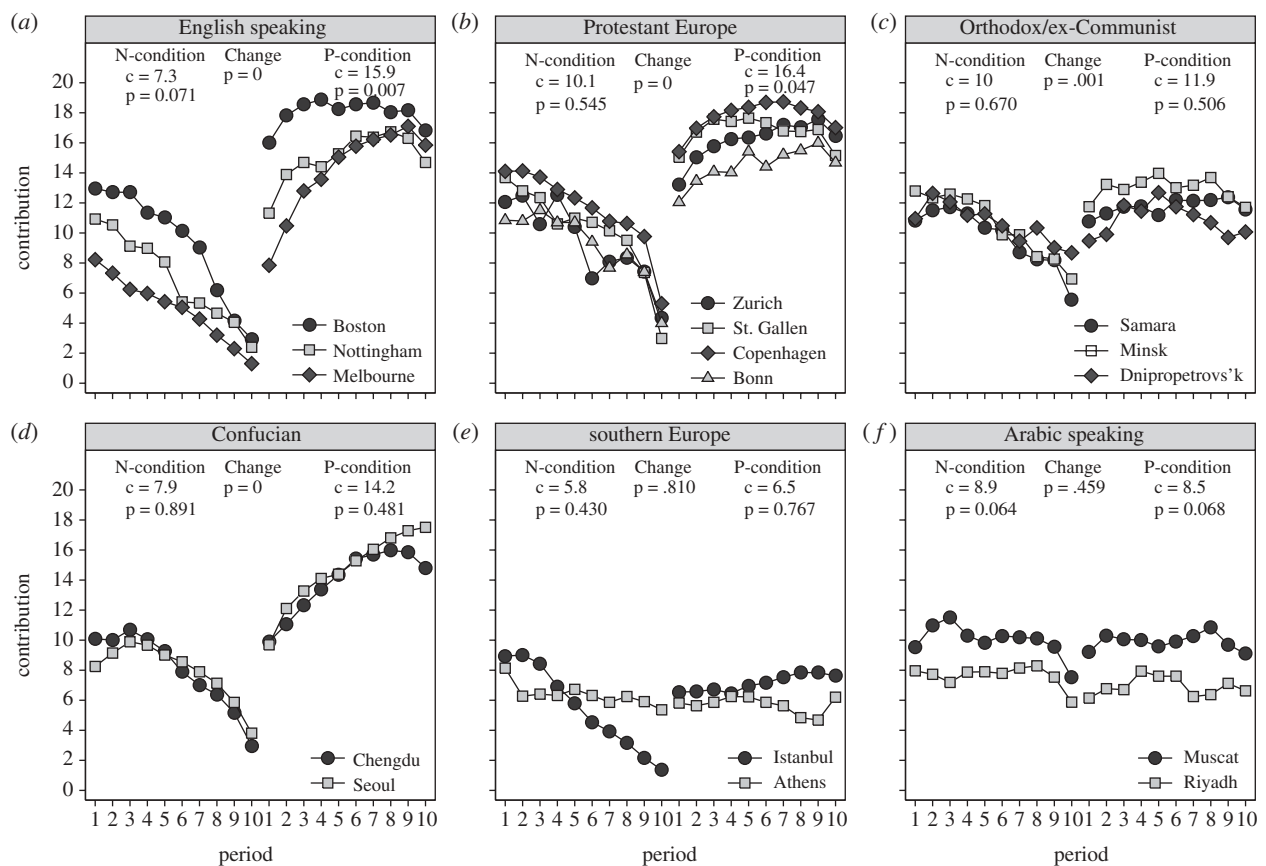


Figure 1. Average contributions in the 16 subject pools during the 10 periods of the N-condition and the P-condition; ‘c’ denotes the average contribution across all periods and subject pools of a given treatment and culture; ‘p’ denotes the p -value of a Kruskal–Wallis test for the equality of contributions of subject pools in a given treatment and culture. ‘Change’ denotes the p -value of a Wilcoxon signed-rank test for the change of contribution between the N-condition and the P-condition. All tests are based on group average contributions over all periods of a respective treatment.

cultural categorization summarized in table 1 and according to treatment condition.

We start with figure 1 and the *N-condition*. The left part of each panel shows the results for the N-condition; ‘c’ indicates the average contribution over the 10 periods. Within all cultures contributions are remarkably similar. According to Kruskal–Wallis tests based on group average contributions across all periods, contributions within a culture are at most weakly significant (in two cultures) and insignificant in four cultures (see p -values indicated in the panels of figure 1). Between cultures, however, contributions are highly significantly different (Kruskal–Wallis test with group averages as independent observations and culture as the grouping variable; $\chi^2(5) = 30.9$, $p = 0.0001$). We interpret this as unambiguous evidence for cultural influences on cooperation in the absence of punishment.

This difference concerns the average level of cooperation. However, all subject pools experience a decline of contributions in the N-condition over time (except subjects in Athens and the two Arabic subject pools, where contributions appear more stable). The explanation of the decline of cooperation is beyond the scope of this paper. We refer the reader to Neugebauer *et al.* (2009) and Fischbacher & Gächter (2010) for analyses of the almost ubiquitous decline of cooperation in finitely repeated public good games. To test whether there are also cultural

differences with regard to the extent of the decline of cooperation, we calculated for each independent group a Spearman rank order correlation of group average contribution and period. We use this correlation coefficient as a test statistic in a Kruskal–Wallis test with the cultural regions as the test groups. We find highly significant differences ($\chi^2(5) = 42.1$, $p = 0.0001$).

We now turn to the analysis of the *P-condition* (illustrated in the right part of each panel). Within a culture the temporal patterns are surprisingly similar. In some of the cultures there is also an indication of significant within-culture variation: cooperation levels are significantly different in two and weakly significantly different in one culture. Across cultures contribution levels are highly significantly different (Kruskal–Wallis test with group averages as independent observations and culture as the grouping variable; $\chi^2(5) = 96.5$, $p = 0.0001$).

Figure 1 (and figures 3 and 4 below) also suggest that there are cultural differences with regard to the *change* of contributions between the N-condition and the P-condition: in four cultures contributions are significantly higher in the P-condition than in the N-condition (with $p < 0.002$) whereas in two cultures this change is not significant (with $p > 0.459$, Wilcoxon signed-ranks tests with group averages as independent observations (see the p -values for ‘change’ indicated in figure 1).

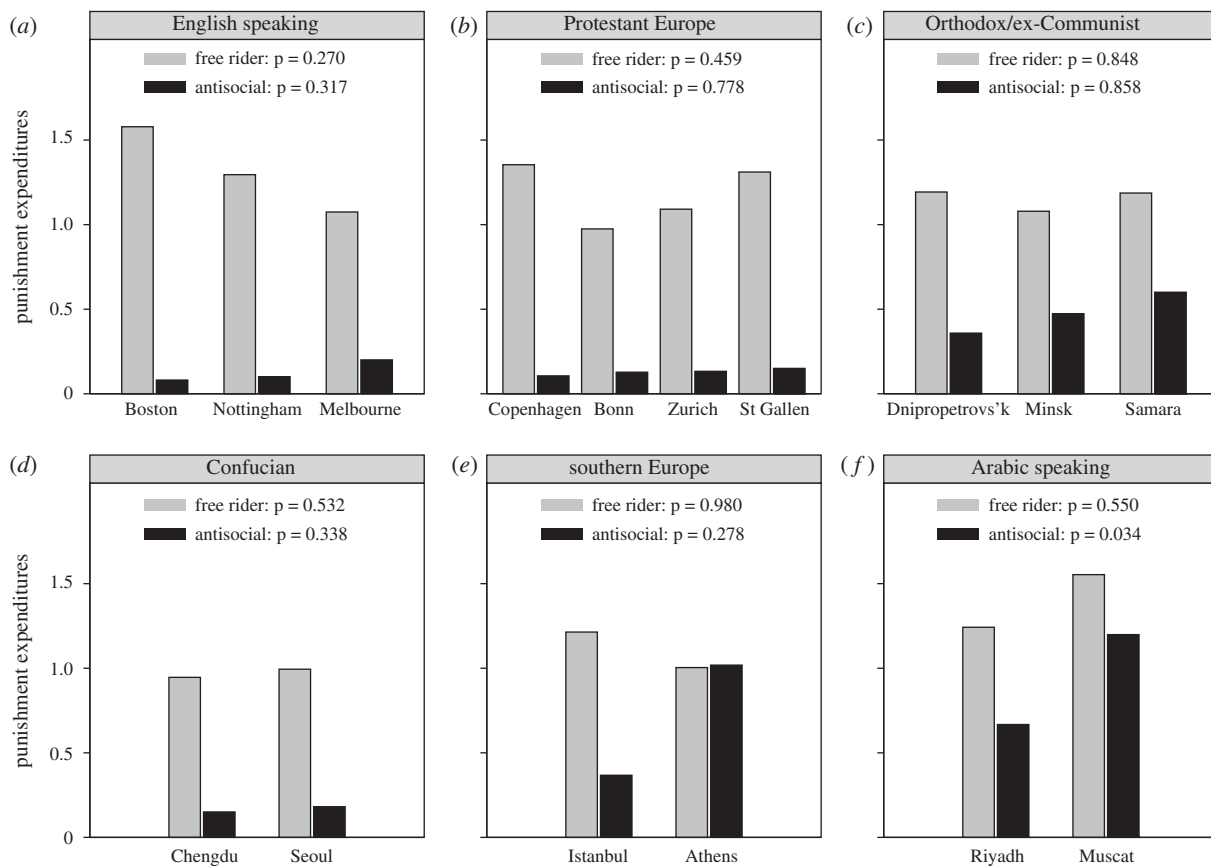


Figure 2. Average expenditures for punishment targeted at subjects with a lower contribution (free-rider punishment) and targeted at subjects with a weakly higher contribution (antisocial punishment) than the punishing subject. p -values are from Kruskal–Wallis tests for differences across subject pools based on the independent group averages.

We conclude from this analysis that there are cultural differences in contributions, in particular in the P-condition. The major part of these cultural differences in the P-condition is most probably owing to differences in punishment. Antisocial behaviour increasingly attracts attention in the study of cooperation (Jensen 2010) but the role of culture remains little explored. Herrmann *et al.* (2008a), table 1, show that contributions are strongly linked to patterns of punishment. In particular they show that contributions in the P-condition depend (i) positively on the initial contribution, (ii) positively on the extent of punishment of free-riding behaviour, and (iii) negatively on antisocial punishment, that is punishment of people who contributed the same or more than the punishing individual. Herrmann *et al.* (2008a) also show that antisocial punishment is strongly linked to norms of civic cooperation in a given society as measured by representative questionnaires in the World Values Survey and the strength of the rule of law in a country (see Herrmann *et al.* (2008b) for further details and references). Both measures differ strongly between the societies of the subject pools of Herrmann *et al.* (2008a). Thus, (antisocial) punishment seems to be linked to the societal background. This observation begs the question of cultural differences in punishment behaviour. Herrmann *et al.* (2008a) have already shown that there are only weakly significant differences in punishment of free-riding behaviour and highly significant differences in antisocial punishment across subject

pools. Are there cultural differences in punishment if we apply our concept of cultural differences?

Figure 2 depicts the extent of average punishment of free-riding behaviour as well as of antisocial punishment per subject pool and grouped for the six cultural areas. Interestingly, with one exception, there are no significant differences in both free rider punishment and antisocial punishment within cultures (based on Kruskal–Wallis tests). Moreover, we find significant differences in punishment across cultures for free rider punishment ($\chi^2(5) = 11.2$, $p = 0.048$) and much stronger cultural differences in antisocial punishment ($\chi^2(5) = 82.5$, $p = 0.0001$).

In addition to the culture-specific changes in contributions between conditions, figures 3 and 4 illustrate two further features of the data, which we shall analyse in more detail in the next step. Figure 3 focuses on the distribution of *individual average contributions* and shows that in the cultures in which punishment leads to a significant behavioural change, the variance of individual contributions is reduced as well. Not very surprisingly, punishment, when it ‘works’, makes people’s contributions more similar (and increases the level of contributions), whereas no such homogenizing effect is visible when punishment is ineffective. In two cultures, the variance of individual contributions even increases in the presence of punishment.

Figure 4 illustrates how *group average contributions* are distributed between conditions and cultures. This is interesting because cooperation in the Herrmann *et al.* (2008a) experiments happened in groups

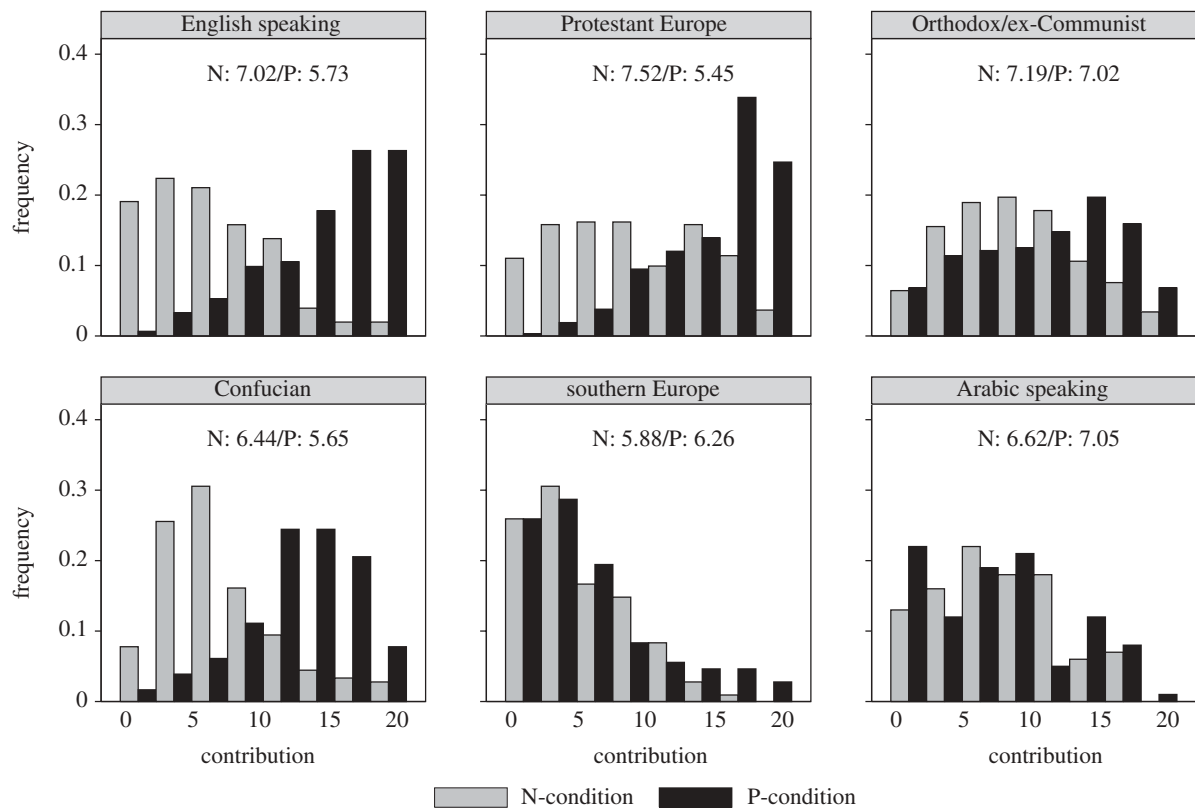


Figure 3. Histograms of individual average contributions in the N- and P-condition for each culture. The numbers in each panel indicate the standard deviation of the contributions in a culture in the two conditions. To measure the standard deviation of contributions independently of the time trend, we calculate a standard deviation for each of the 10 periods and report the average standard deviation across the 10 periods.

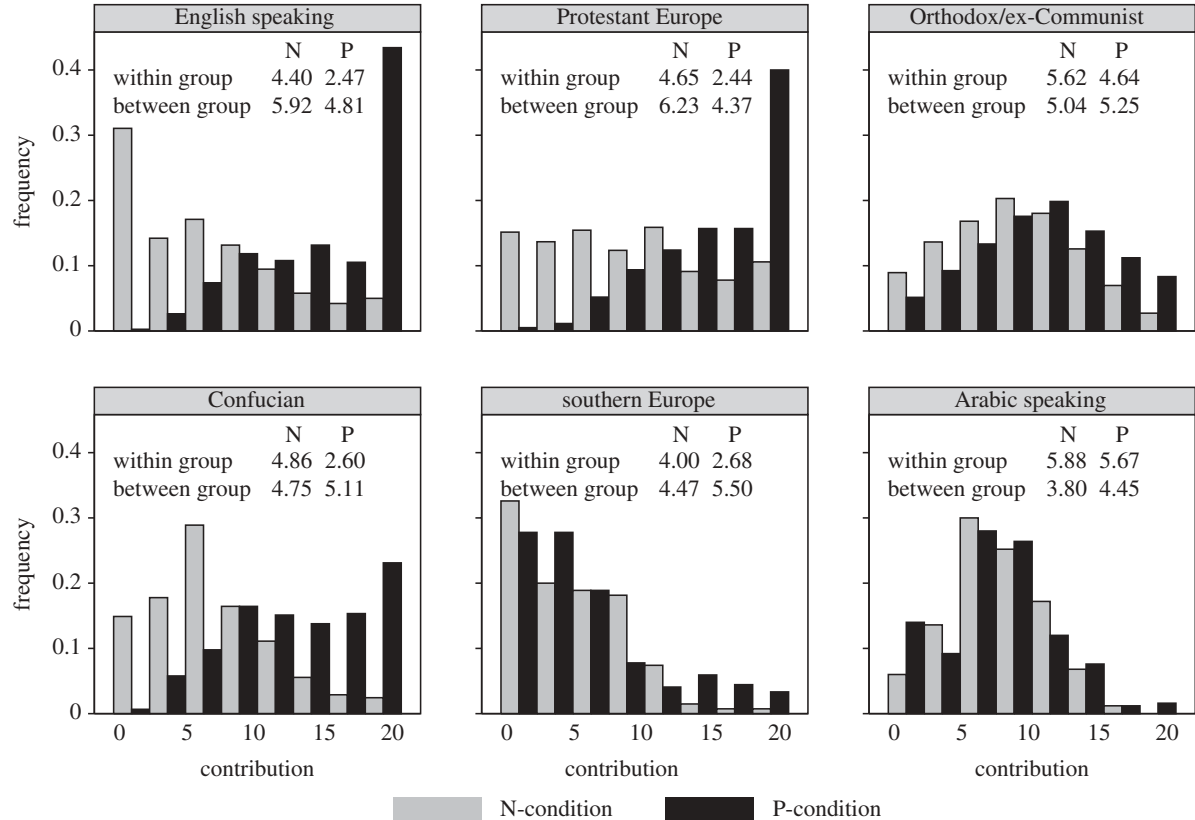


Figure 4. Histograms of average group contributions per period in the N- and P-condition for each culture. The numbers indicate standard deviations of the contributions within a group and between groups. For the within-group standard deviations, we calculate the standard deviations of the four contributions in a group in each period and average over all periods and groups within a culture. For the between-group measure, we calculate the standard deviation of all group averages within a culture and a period. The numbers show the average over the 10 periods for the N- and the P-condition.

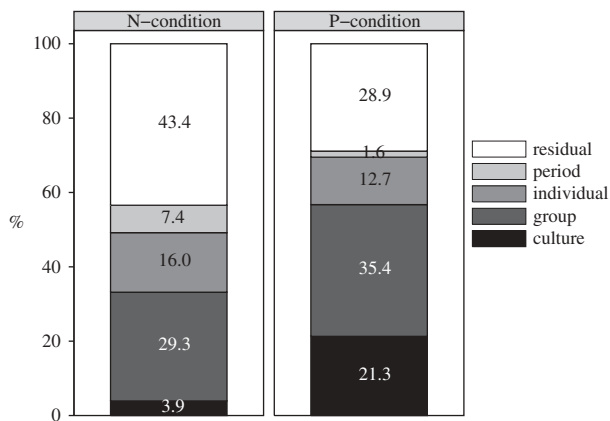


Figure 5. Decomposition of the coefficient of determination for contributions in the two treatment conditions.

with fixed memberships over time and groups might have been 'locked' into a particular path-dependent contribution pattern, for example owing to a frequent tendency of conditional cooperation (e.g. Gächter & Thöni 2005; Kurzban & Houser 2005; Gunnthorsdottir *et al.* 2007; Fischbacher & Gächter 2010). Such path-dependency might lead to substantially different group average contributions, and therefore to large between-group variance. Moreover, the presence of punishment might affect both the between-group variance (by making groups more homogeneous) and the within-group variance. We find that the introduction of punishment reduces the within-group variance in all six cultures. The effect on the between-group variance is more diverse: in four of the six cultures the between-group variance increases and in two cultures it decreases.

We conclude from this descriptive analysis that cultural differences in contribution decisions exist without doubt. In our next step, we are interested in the relative fraction of the variance that is due to individuals, groups and in particular culture in contributions in both the N- and the P-conditions. For this purpose, we use the nested ANOVA model described in §2 to decompose the explanatory power of our measure for culture, group composition and individual differences.

Figure 5 shows the R^2 -associated to our explanatory variables for the N- and the P-conditions. It measures the sum of squares associated with the explanatory variable divided by the total sum of squares in the contribution decisions. Bar heights depict the fraction of the variance that is explained by the corresponding variable. The lowest part of a bar depicts the fraction of the variance explained by Culture. In the N-condition, the cultural variation in our subject pool explains only a small amount of the variance (3.9%). Group-level differences (that is, between-group variance) account for additional 29.3 per cent of the variation in contributions, and a further 16.0 per cent can be explained by individual-fixed effects. Time effects account for 7.4 per cent of the variation. Finally, 43.4 per cent of the variation remains unexplained by our model.

Comparing the results of the N-condition to the results of the P-condition reveals striking differences.

First, a much smaller portion of the variance in contributions remains unexplained. Individual and period effects lose much of their explanatory power while Group and Culture gain in importance. In particular, the percentage of the variance explained by our cultural classification is more than five times larger in the P-condition than in the N-condition.

Are these fractions of explained variance large? This is an important question, because even in the absence of any systematic cultural, group or individual effects the ANOVA model would provide some non-zero R^2 . We ran 100 ANOVAs with simulated contributions (all contributions in 0,1,...,20 were drawn with equal probability). The explanatory power of Culture in the absence of systematic cultural variation is very close to zero (mean: 0.043%, s.d. 0.025). Consequently, the influence of Culture is far beyond the effect that would show up in the absence of cultural variation. The same is true for Group and Individual effects, as well as Period effects.⁶

4. SUMMARY AND CONCLUDING REMARKS

In this paper, we have analysed an experimental dataset by Herrmann *et al.* (2008a) who ran comparable public good experiments with and without punishment in 16 subject pools from six distinct cultural areas around the world. This dataset allows us to show that cultural differences in cooperation exist in the sense that within-cultural variation is smaller than the between-cultural variation. Moreover, we found that for the extent of cooperation we observe, cultural variation is a particularly important source of variation in the presence of punishment opportunities. This is due to large cultural differences in punishment. In the absence of punishment, individual ('micro-level') variation is much more important than cultural ('macro-level') variation, whereas the opposite is true in the presence of punishment. Group-level differences (the 'meso-level') are very important both in the presence and the absence of punishment.

We know from numerous experiments that individual heterogeneity is an important source of variation that can translate into important aggregate-level differences in outcomes (Camerer & Fehr 2006; Gächter & Thöni *in press*). Our dataset confirms this insight by showing that individual variation and group-level variation are both important sources of the overall variation. The importance of our finding of culture effects in addition to individual-level and group-level differences is that, holding everything else constant, differences in cultural background can lead to differences in behaviour in otherwise identical environments. Thus, accounting for individual and implied group-level differences is not enough to understand the whole breadth of variation in cooperation. Culture needs to be accounted for.

We conclude with two caveats and future research questions. First, in this analysis, we have only demonstrated the existence and quantitative importance of cultural differences. Our approach cannot explain where the differences come from. Herrmann *et al.* (2008a) found large differences in cooperation only in the presence of punishment and owing to large

differences in punishment across subject pools. Ascertaining why these cultural differences in punishment occur is an interesting task for future research. Second, we have drawn our conclusions from comparing subjects who are very similar with regard to their socio-economic status and other socio-demographic characteristics. However, in every society there exist various social groups who might also show a large variation in cooperative behaviour (e.g. Ockenfels & Weimann 1999; Fehr et al. 2002a; Bellemare & Kröger 2007; Hong & Bohnet 2007; Hoff et al. 2009; Kocher et al. 2009; Gächter & Herrmann in press; Henrich et al. in press). It is an important task for future research to understand this sort of variation relative to the sources of variation we have identified in this paper.

We gratefully acknowledge financial support from the University of Nottingham, the Latsis Foundation (Geneva), and the EU-TMR Research Network ENDEAR (FMRX-CT98-0238). We received helpful comments from the editors and referees, as well as various workshop audiences, in particular the Arts and Humanities Research Council workshops Culture and the Mind in Sheffield, and from Peter Egger and Conny Wunsch. This paper is part of the MacArthur Foundation Network on Economic Environments and the Evolution of Individual Preferences and Social Norms.

ENDNOTES

¹Gächter et al. (2004) and Thöni et al. (2009) show that, on the individual level and within a given culture, there is a connection between questionnaire items as used in the World Values Survey and cooperation in public goods games. Such a relationship has also been established in trust games, which also contain an element of cooperation (Ermisch et al. 2009).

²See Friedman & Sunder (1994) for an introduction to methods in experimental economics; Guala (2005), Falk & Heckman (2009), Bardsley et al. (2010), Croson & Gächter (2010) and Smith (2010) for a discussion of the methodology of experimental economics. Gächter & Herrmann (2009) provide an overview of experiments on cooperation and punishment.

³The most important experimental tool in these studies is the ultimatum game (Güth et al. 1982). For a comprehensive analysis and cross-cultural comparison of ultimatum bargaining games see Oosterbeek et al. (2004).

⁴Gächter & Herrmann (2009) applied this methodology to one-shot experiments conducted with students ($n = 606$) in two Swiss subject pools and two Russian subject pools. According to several measures, Russia and Switzerland are culturally very distinct societies. The results show within-cultural similarity but strong between-cultural differences.

⁵In general, the ANOVA does not allow for an unambiguous disaggregation of the coefficient of determination. In our analysis this is possible because all exogenous variables are orthogonal and our sample is balanced.

⁶The simulation ($n = 100$) for Group yields mean = 2.42%, s.d. = 0.197; for Individual mean = 7.51%, s.d. = 0.330. Period: mean = 0.082%, s.d. = 0.039.

REFERENCES

- Axelrod, R. 1984 *The evolution of cooperation*. New York, NY: Basic Books.
- Bardsley, N. & Moffatt, P. G. 2007 The experimetrics of public goods: inferring motivations from contributions. *Theory Decis.* **62**, 161–193. (doi:10.1007/s11238-006-9013-3)
- Bardsley, N., Cubitt, R., Loomes, G., Moffatt, P., Starmer, C. & Sugden, R. 2010 *Experimental economics: rethinking the rules*. Princeton, NJ: Princeton University Press.
- Bednar, J. & Page, S. 2007 Can game(s) theory explain culture? The emergence of cultural behavior within multiple Games. *Ration. Soc.* **19**, 65–97. (doi:10.1177/1043463107075108)
- Bellemare, C. & Kröger, S. 2007 On representative social capital. *Eur. Econ. Rev.* **51**, 183–202. (doi:10.1016/j.euroecorev.2006.03.006)
- Bergmüller, R., Schürch, R. & Hamilton, I. M. 2010 Evolutionary causes and consequences of consistent individual variation in cooperative behaviour. *Phil. Trans. R. Soc. B* **365**, 2751–2764. (doi:10.1098/rstb.2010.0124)
- 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. (doi:10.1257/aer.98.1.294)
- Bohnet, I., Herrmann, B. & Zeckhauser, R. 2010 Trust and reference points for trustworthiness in Gulf and Western countries. *Quart. J. Econ.* **125**. (doi:10.1162/qjec.2010.125.2.811)
- Bornhorst, F., Ichino, A., Kirchkamp, O., Schlag, K. H. & Winter, E. In press. Similarities and differences when building trust: the role of cultures. *Exp. Econ.* (doi:10.1007/s10683-010-9240-x)
- Bowles, S. 1998 Endogenous preferences: the cultural consequences of markets and other economic institutions. *J. Econ. Lit.* **36**, 75–111.
- Boyd, R. & Richerson, P. J. 2005 *The origin and evolution of cultures*. In *Evolution and cognitions*. Oxford, UK: Oxford University Press.
- Buchan, N., Croson, R. & Dawes, R. M. 2002 Swift neighbors and persistent strangers: a cross-cultural investigation of trust and reciprocity in social exchange. *Am. Sociol. Rev.* **108**, 168–206. (doi:10.1086/344546)
- Buchan, N. R., Grimalda, G., Wilson, R. K., Brewer, M., Fatas, E. & Foddy, M. 2009 Globalization and human cooperation. *Proc. Natl Acad. Sci. USA* **106**, 4138–4142. (doi:10.1073/pnas.0809522106)
- Camerer, C. F. & Fehr, E. 2006 When does ‘economic man’ dominate social behavior? *Science* **311**, 47–52. (doi:10.1126/science.1110600)
- Chuah, S.-H., Hoffmann, R., Jones, M. & Williams, G. 2007 Do cultures clash? Evidence from cross-national ultimatum game experiments. *J. Econ. Behav. Organ.* **64**, 35–48. (doi:10.1016/j.jebo.2006.04.006)
- Chuah, S.-H., Hoffmann, R., Jones, M. & Williams, G. 2009 An economic anatomy of culture: attitudes and behaviour in inter- and intra-national ultimatum game experiments. *J. Econ. Psychol.* **30**, 732–744. (doi:10.1016/j.joep.2009.06.004)
- Cohen, D. 2001 Cultural variation: considerations and implications. *Psychol. Bull.* **127**, 451–471. (doi:10.1037/0033-2909.127.4.451)
- Croson, R. 2007 Theories of commitment, altruism and reciprocity: evidence from linear public goods games. *Econ. Inquiry* **45**, 199–216. (doi:10.1111/j.1465-7295.2006.00006.x)
- Croson, R. & Gächter, S. 2010 The science of experimental economics. *J. Econ. Behav. Organ.* **73**, 122–131.
- Dufwenberg, M., Gächter, S. & Hennig-Schmidt, H. 2006 *The framing of games and the psychology of strategic choice*. Nottingham, UK: University of Nottingham. (CeDEx Discussion Paper 2006–2020).
- Ermisch, J., Gambetta, D., Laurie, H., Siedler, T. & Noah Uhrig, S. C. 2009 Measuring people’s trust. *J. R. Stat. Soc. Ser. A (Statistics in Society)* **172**, 749–769. (doi:10.1111/j.1467-985X.2009.00591.x)
- Falk, A. & Heckman, J. J. 2009 Lab experiments are a major source of knowledge in the social sciences. *Science* **326**, 535–538. (doi:10.1126/science.1168244)

- Fehr, E. & Fischbacher, U. 2004 Social norms and human cooperation. *Trends Cogn. Sci.* **8**, 185–190. (doi:10.1016/j.tics.2004.02.007)
- Fehr, E. & Gächter, S. 2000 Cooperation and punishment in public goods experiments. *Am. Econ. Rev.* **90**, 980–994.
- Fehr, E. & Gächter, S. 2002 Altruistic punishment in humans. *Nature* **415**, 137–140. (doi:10.1038/415137a)
- Fehr, E. & Schmidt, K. M. 2006 The economics of fairness, reciprocity and altruism—experimental evidence and new theories. In *Handbook of the economics of giving, altruism and reciprocity*, vol. 1 (eds S.-C. Kolm & J. M. Ythier), pp. 615–691. Amsterdam, The Netherlands: Elsevier.
- Fehr, E., Fischbacher, U., von Rosenbladt, B., Schupp, J. & Wagner, G. G. 2002a A nationwide laboratory. Examining trust and trustworthiness by integrating behavioral experiments into representative surveys. *Schmoller's Jahrbuch* **122**, 519–542.
- Fehr, E., Gächter, S. & Fischbacher, U. 2002b Strong reciprocity, human cooperation, and the enforcement of social norms. *Hum. Nat.* **13**, 1–25. (doi:10.1007/s12110-002-1012-7)
- Fernández, R. 2007 Women, work, and culture. *J. Eur. Econ. Assoc.* **5**, 305–332. (doi:10.1162/jeea.2007.5.2-3.305)
- Fernández, R. 2008 Culture and economics. In *The new Palgrave dictionary of economics online*. Basingstoke, UK: Palgrave Macmillan. See http://www.dictionaryofeconomics.com/article?id=pde2008_E000282 (doi:10.1057/9780230226203.0346)
- Fischbacher, U. & Gächter, S. 2010 Social preferences, beliefs, and the dynamics of free riding in public good experiments. *Am. Econ. Rev.* **100**, 541–556. (doi:10.1257/aer.100.1.541)
- Fischbacher, U., Gächter, S. & Fehr, E. 2001 Are people conditionally cooperative? Evidence from a public goods experiment. *Econ. Lett.* **71**, 397–404. (doi:10.1016/S0165-1765(01)00394-9)
- Friedman, D. & Sunder, S. 1994 *Experimental methods. a primer for economists*. Cambridge, UK: Cambridge University Press.
- Fudenberg, D. & Maskin, E. 1986 The folk theorem in repeated games with discounting or with incomplete information. *Econometrica* **54**, 533–556. (doi:10.2307/1911307)
- Gächter, S. 2007 Conditional cooperation: behavioral regularities from the lab and the field and their policy implications. In *Psychology and economics: a promising new cross-disciplinary field* (CESifo Seminar Series) (eds B. S. Frey & A. Stutzer), pp. 19–50. Cambridge, MA: MIT Press.
- Gächter, S. & Herrmann, B. 2009 Reciprocity, culture, and human cooperation: previous insights and a new cross-cultural experiment. *Phil. Trans. R. Soc. B* **364**, 791–806. (doi:10.1098/rstb.2008.0275)
- Gächter, S. & Herrmann, B. In press. The limits of self-governance when cooperators get punished—experimental evidence from urban and rural Russia. *Eur. Econ. Rev.* (doi:10.1016/j.euroecorev.2010.04.003)
- Gächter, S. & Thöni, C. 2005 Social learning and voluntary cooperation among like-minded people. *J. Eur. Econ. Assoc.* **3**, 303–314.
- Gächter, S. & Thöni, C. In press. Micromotives, microstructure and macrobehavior: the case of voluntary cooperation. *J. Math. Sociol.*
- 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. (doi:10.1016/j.jebo.2003.11.006)
- Gächter, S., Herrmann, B. & Thöni, C. 2005 Cross-cultural differences in norm enforcement. *Behav. Brain Sci.* **28**, 822.
- Gintis, H. In press. Gene-culture coevolution and the nature of human sociality. *Phil. Trans. R. Soc. B*.
- Gintis, H., Bowles, S., Boyd, R. & Fehr, E. (eds) 2005 *Moral sentiments and material interests. The foundations of cooperation in economic life*. Cambridge, MA: MIT Press.
- Greif, A. 1994 Cultural beliefs and the organization of society: a historical reflection on collectivist and individualist societies. *J. Political Econ.* **102**, 912–950. (doi:10.1086/261959)
- Guala, F. 2005 *The methodology of experimental economics*. Cambridge, UK: Cambridge University Press.
- Guiso, L., Sapienza, P. & Zingales, L. 2006 Does culture affect economic outcomes? *J. Econ. Perspect.* **20**, 23–48. (doi:10.1257/jep.20.2.23)
- Guiso, L., Sapienza, P. & Zingales, L. 2008 Social capital as good culture. *J. Eur. Econ. Assoc.* **6**, 295–320. (doi:10.1162/JEEA.2008.6.2-3.295)
- Gunnthorsdottir, A., Houser, D. & McCabe, K. 2007 Disposition, history and contributions in public goods experiments. *J. Behav. Organ.* **62**, 304–315. (doi:10.1016/j.jebo.2005.03.008)
- Güth, W., Schmittberger, R. & Schwarze, B. 1982 An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* **3**, 367–388. (doi:10.1016/0167-2681(82)90011-7)
- Hayashi, N., Ostrom, E., Walker, J. & Yamagishi, T. 1999 Reciprocity, trust, and the sense of control—a cross-societal study. *Ration. Soc.* **11**, 27–46. (doi:10.1177/104346399011001002)
- Heine, S. J. & Buchtel, E. E. 2009 Personality: the universal and the culturally specific. *Annu. Rev. Psychol.* **60**, 369–394. (doi:10.1146/annurev.psych.60.110707.163655)
- Heine, S. J. & Ruby, M. B. 2010 Cultural psychology. *Wiley Interdiscip. Rev. Cogn. Sci.* **1**, 254–266.
- Henrich, J. 2004 Cultural group selection, coevolutionary processes and large-scale cooperation. *J. Econ. Behav. Organ.* **53**, 3–35. (doi:10.1016/S0167-2681(03)00094-5)
- Henrich, J. & Henrich, N. 2007 Why humans cooperate: a cultural and evolutionary explanation. *Evolution and cognition series*. Oxford, UK: Oxford University Press.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C. F., Fehr, E., Gintis, H. & McElreath, R. 2001 In search of *Homo economicus*: behavioral experiments in 15 small-scale societies. *Am. Econ. Rev.* **91**, 73–78.
- Henrich, J. *et al.* 2005 ‘Economic man’ in cross-cultural perspective: behavioral experiments in 15 small-scale societies. *Behav. Brain Sci.* **28**, 795–855.
- Henrich, J. *et al.* 2006 Costly punishment across human societies. *Science* **312**, 1767–1770. (doi:10.1126/science.1127333)
- Henrich, J. *et al.* 2010 Markets, religion, community size, and the evolution of fairness and punishment. *Science* **327**, 1480–1484. (doi:10.1126/science.1182238)
- Henrich, J., Heine, S. J. & Noranyazan, A. In press. The weirdest people in the world? *Behav. Brain Sci.*
- Herrmann, B. & Thöni, C. 2009 Measuring conditional cooperation: a replication study in Russia. *Exp. Econ.* **12**, 87–92. (doi:10.1007/s10683-008-9197-1)
- Herrmann, E., Call, J., Hernandez-Lloreda, M. V., Hare, B. & Tomasello, M. 2007 Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science* **317**, 1360–1366. (doi:10.1126/science.1146282)
- Herrmann, B., Thöni, C. & Gächter, S. 2008a Antisocial punishment across societies. *Science* **319**, 1362–1367. (doi:10.1126/science.1153808)
- Herrmann, B., Thöni, C. & Gächter, S. 2008b Supporting online material for antisocial punishment across societies. *Science* **319**. See <http://www.sciencemag.org/cgi/data/319/5868/1362/DC1/1>. (doi:10.1126/science.1153808)

- Hoff, K., Kshetramade, M. & Fehr, E. 2009 *Caste and punishment: the legacy of caste culture in norm enforcement*. Bonn, Germany: Institute for the Future of Labor. (IZA Discussion Paper No. 4350).
- Hofstede, G. 2001 *Culture's consequences: comparing values, behaviors, institutions, and organizations across nations*. Thousand Oaks, CA: Sage.
- Holm, H. J. & Danielson, A. 2005 Tropic trust versus Nordic trust: experimental evidence from Tanzania and Sweden. *Econ. J.* **115**, 105–132.
- Hong, K. & Bohnet, I. 2007 Status and distrust: the relevance of inequality and betrayal aversion. *J. Econ. Psychol.* **28**, 197–213. (doi:10.1016/j.joep.2006.06.003)
- Inglehart, R. 1997 *Modernization and postmodernization: cultural, economic and political change in 43 societies*. Princeton, NJ: Princeton University Press.
- Inglehart, R. & Baker, W. E. 2000 Modernization, cultural change, and the persistence of traditional values. *Am. Sociol. Rev.* **65**, 19–51. (doi:10.2307/2657288)
- Jensen, K. 2010 Punishment and spite, the dark side of cooperation. *Phil. Trans. R. Soc. B* **365**, 2635–2650. (doi:10.1098/rstb.2010.0146)
- Kachelmeier, S. J. & Shehata, M. 1997 Internal auditing and voluntary cooperation in firms: a cross-cultural experiment. *Account. Rev.* **72**, 407–431.
- Kocher, M. G., Cherry, T., Kroll, S., Netzer, R. J. & Sutter, M. 2008 Conditional cooperation on three continents. *Econ. Lett.* **101**, 175–178. (doi:10.1016/j.econlet.2008.07.015)
- Kocher, M., Martinsson, P. & Visser, M. 2009 *Social background, cooperative behavior, and norm enforcement*. Gothenburg, Sweden: University of Gothenburg. (Working Papers In Economics No. 385).
- Kurzban, R. & Houser, D. 2005 Experiments investigating cooperative types in humans: a complement to evolutionary theory and simulations. *Proc. Natl Acad. Sci. USA* **102**, 1803–1807. (doi:10.1073/pnas.0408759102)
- McNamara, J. M. & Leimar, O. 2010 Variation and the response to variation as a basis for successful cooperation. *Phil. Trans. R. Soc. B* **365**, 2627–2633. (doi:10.1098/rstb.2010.0159)
- Markus, H. & Kitayama, S. 1991 Culture and the self. Implications for cognition, emotion, and motivation. *Psychol. Rev.* **98**, 224–253. (doi:10.1037/0033-295X.98.2.224)
- Muller, L., Sefton, M., Steinberg, R. & Vesterlund, L. 2008 Strategic behavior and learning in repeated voluntary-contribution experiments. *J. Econ. Behav. Organ.* **67**, 782–793. (doi:10.1016/j.jebo.2007.09.001)
- Nettle, D. 2009 Beyond nature versus culture: cultural variation as an evolved characteristic. *J. R. Anthropol. Inst.* **15**, 223–240. (doi:10.1111/j.1467-9655.2009.01561.x)
- Neugebauer, T., Perote, J., Schmidt, U. & Loos, M. 2009 Self-biased conditional cooperation: on the decline of cooperation in repeated public goods experiments. *J. Econ. Psychol.* **30**, 52–60. (doi:10.1016/j.joep.2008.04.005)
- Nisbett, R. 2003 *The geography of thought: how Asians and Westerners think differently and why*. New York, NY: Free Press.
- Nisbett, R. & Cohen, D. 1996 *Culture of honor: the psychology of violence in the South*. Boulder, CO: Westview Press.
- Nowak, M. A. 2006 Five rules for the evolution of cooperation. *Science* **314**, 1560–1563. (doi:10.1126/science.1133755)
- Ockenfels, A. & Weimann, J. 1999 Types and patterns: an experimental East-West-German comparison of cooperation and solidarity. *J. Public Econ.* **71**, 275–287. (doi:10.1016/S0047-2727(98)00072-3)
- Oosterbeek, H., Sloof, R. & van de Kuilen, G. 2004 Cultural differences in ultimatum game experiments: evidence from a meta-analysis. *Exp. Econ.* **7**, 171–188. (doi:10.1023/B:EXEC.0000026978.14316.74)
- Ostrom, E., Walker, J. M. & Gardner, R. 1992 Covenants with and without a sword—self-governance is possible. *Am. Political Sci. Rev.* **86**, 404–417.
- Rendell, L. et al. 2010 Why copy others? Insights from the social learning strategies tournament. *Science* **328**, 208–213. (doi:10.1126/science.1184719)
- Richerson, P. J. & Boyd, R. 1999 Complex societies—the evolutionary origins of a crude superorganism. *Hum. Nat.* **10**, 253–289. (doi:10.1007/s12110-999-1004-y)
- 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.
- Sigmund, K. 2010 *The calculus of selfishness*. Princeton, NJ: Princeton University Press.
- Smith, V. L. 2010 Theory and experiment: what are the questions? *J. Econ. Behav. Organ.* **73**, 3–15. (doi:10.1016/j.jebo.2009.02.008)
- Sober, E. & Wilson, D. S. 1998 *The evolution and psychology of unselfish behavior*. Cambridge, MA: Harvard University Press.
- Tabellini, G. 2008a Institutions and culture. *J. Eur. Econ. Assoc.* **6**, 255–294. (doi:10.1162/JEEA.2008.6.2-3.255)
- Tabellini, G. 2008b The scope of cooperation: values and incentives. *Quart. J. Econ.* **123**, 905–950. (doi:10.1162/qjec.2008.123.3.905)
- Thöni, C., Tyran, J.-R. & Wengström, E. 2009 *Micro-foundations of social capital*. Copenhagen, Denmark: Department of Economics, University of Copenhagen. (Discussion Paper No. 09-24).
- Tomasello, M. 2009 *Why we cooperate*. Cambridge, MA: MIT Press.
- Tomasello, M., Carpenter, M., Call, J., Behne, T. & Moll, H. 2005 Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* **28**, 675–691.
- Wu, J. J., Zhang, B. Y., Zhou, Z. X., He, Q. Q., Zheng, X. D., Cressman, R. & Tao, Y. 2009 Costly punishment does not always increase cooperation. *Proc. Natl Acad. Sci. USA* **106**, 17 448–17 451. (doi:10.1073/pnas.0905918106)
- Yamagishi, T. 1988 The provision of a sanctioning system in the United-States and Japan. *Social Psychol. Quart.* **51**, 265–271. (doi:10.2307/2786924)
- Yamagishi, T. & Yamagishi, M. 1994 Trust and commitment in the United-States and Japan. *Motiv. Emotion* **18**, 129–166. (doi:10.1007/BF02249397)
- Yamagishi, T., Cook, K. S. & Watabe, M. 1998 Uncertainty, trust, and commitment formation in the United States and Japan. *Am. J. Sociol.* **104**, 165–194. (doi:10.1086/210005)