Table 1 (Gächter et al.). *Main results from cross-cultural experiments on cooperation and punishment*

	Mean contribution in treatment				
	No. subject	N	P	Mean punishment of free riders	Mean punishment of cooperators
Zurich	140	8.5	16.2	1.22	0.15
Strasbourg	96	8.0	11.3	0.86	0.34
Minsk	68	10.5	12.9	1.11	0.51
Samara	152	10.4	11.5	1.15	0.64
Kruskal-Wallis tests	—	0.10	0.00	0.37	0.00

voluntary contributions to public goods (e.g., Brandts et al. 2004; Kachelmeier & Shehata 1997;). Thus, many more systematic cross-cultural experiments would be needed before the lack of cross-cultural variation is an established fact in games other than the ultimatum game.

Second, and this is our main point, if we move beyond "basic" games, and look at norm enforcement, differences between social groups are likely to emerge even if AMI and PC are similar. The basis for this belief is experiments on public goods games with punishment, which we see as a model of norm enforcement. We (Gächter et al., in preparation) ran a standard linear public goods game, very similar to the one used by Fehr and Gächter (2000a). We conducted the experiments in Zürich (Switzerland), Strasbourg (France), Minsk (Belarus), and Samara (Russia). Participants (undergraduates from the respective universities at an average age of 20) were divided into groups of four members who played the game in the same group for ten periods. In the non-punishment condition (the "N-treatment"), subjects had to decide simultaneously on their contribution to a public good. In our terminology, this game may classify as a "basic game," because cooperation is the only issue. In the punishment condition (the "P-treatment"), a second stage was added where each subject could punish each group member at its own cost. One punishment unit cost the punishing subject one money unit and reduced the punished subject's payoff from the first stage by three money units. We applied standard methods to ensure cross-cultural comparability (e.g., instructions were translated into Russian or French, and translated back into German to control for language-induced differences in meaning, etc.).

Table 1 presents the key results. We report both mean contributions over all periods in the N and in the P treatments. In the N-treatment we find only minor differences in cooperation rates between our four subject pools. The differences are not statistically significant. This finding is consistent with (1) comparable public goods experiments (Brandts et al. 2004; Kachelmeier & Shehata 1997), and (2) with the hypothesis that cross-cultural differences are small in basic games.

Yet, with the introduction of the opportunity to punish each other, strong differences emerge: Compared to their average contribution in the N-treatment, the Swiss students increase their contributions by 90 percent, while the French subjects increase their contributions by 41 percent. Belarusian and Russian students increase their contribution only by 23 percent and 11 percent, respectively. The increase is significant at the 5 percent level only for the Swiss subjects.

The key to understanding this result is punishment behaviour. Table 1 shows that the four subject pools differ greatly with respect to how they punish "free riders" (defined as group members who contributed less than the punishing subject) and "cooperators" (group members who contributed at least as much as the punisher). For instance, the Zurich subjects punish a "free rider" on average by 1.22 points and a "cooperator" by 0.15 points. The Strasbourg subjects contribute very similar amounts as the subjects in Zurich in the N treatment but reach substantially lower contribution levels in the P treatment. At the same time, their punishment is much less clearly directed towards the free riders. The comparison with Zurich suggests that differences with respect to punishment behaviour may occur even in social groups of quite similar cultural proximity (Strasbourg and Zurich are less than 140 miles apart). The Minsk and Samara subjects punish free riders similarly as do the Zurich subjects, but punish cooperators roughly four times as harshly as the Zurich subjects. Further experiments and data analyses suggest that much of the punishment of cooperators is punishment by free riders in revenge of the punishment the free riders anticipated to receive from the cooperators.

A further data analysis reveals that punishment can successfully solve the free rider problem only when (1) people predominantly punish the free riders sufficiently strongly; (2) the free riders therefore increase their contributions to avoid punishment; and (3) cooperators do not get punished. The experiments show that there are strong differences between groups with respect to the validity of these conditions. This holds despite a very similar readiness to cooperate in the absence of punishment. Punishment is not only about inflicting material sanctions; it also expresses a normative view about unacceptable behaviour. Punishment is also emotion-laden and may trigger revengeful feelings and/or defiance in the punished subject. Both the normative and emotive perception may differ strongly even between sociologically rather uniform subject pools. Once we move away from "basic games," we might uncover surprising and substantial behavioural differences even between student subject pools.

Is the Ultimatum Game a three-body affair?

Gerd Gigerenzer and Thalia Gigerenzer

Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, 14195 Berlin, Germany. **gigerenzer@mpib-berlin.mpg.de**
http://ntfm.mpib-berlin.mpg.de/mpib/FMPro

Abstract: The Ultimatum Game is commonly interpreted as a two-person bargaining game. The third person who donates and may withdraw the money is not included in the theoretical equations, but treated like a neutral measurement instrument. Yet in a cross-cultural analysis it seems necessary to consider the possibility that the thoughts of a player – strategic, altruistic, selfish, or concerned about reputation – are influenced by both an anonymous second player *and* the non-anonymous experimenter.

The behavior of people in the Ultimatum Game (UG) has been analyzed in terms of a two-person interaction between a proposer

https://doi.org/10.1017/S0140525X05300148 Published online by Cambridge University Press

and a responder. Yet there is a third person involved: the experimenter who donates the money and has the power to withdraw it if the responder does not accept the proposer's offer. Since the third person has no name in the theoretical analyses of the UG, let us call him or her the donor. Whether or not the donor can be treated as a neutral observer, equivalent to a measurement instrument, seems to be particularly interesting in cross-cultural comparisons. When an experimenter walks into one of the 15 small-scale societies, he or she represents a technologically advanced tribe and is likely to stand out more than when in a university lab. Both the proposer (she) and the responder (he) know that the donor will record their choices, and they might not be indifferent to the impression their behavior has on the donor. In addition, the responder may realize that by accepting he might take money away from the donor, whereas if he rejects the offer, he will give money to the donor. This three-body perspective differs from the theoretical treatment of the donor as a neutral figure, whose only task is to explain the rules and record the behavior. Our question is: Should we ignore the third person in a cross-cultural study of the Ultimatum Game?

Like Henrich et al.'s abundant results, our commentary poses more questions than it provides answers. Yet there are good reasons to consider the possibility that the behavior of the proposer is not simply a function of his expectations about the responder, or of some stable social preferences, but is also targeted at the donor. The UG is supposedly played anonymously, a term that describes the relation between proposer and responder, whereas in fact there is no anonymity between the two and the donor. In Henrich et al.'s analysis, a choice between a selfish or altruistic offer is assumed to reflect the proposer's social preferences or expectations concerning the responder. Yet, seen as a three-body game, his or her choice could reflect her goals and expectations concerning the donor as well. A proposer might want to appear generous instead of greedy in the eyes of the donor (who is not anonymous) rather than before the responder (who is anonymous). A proposer may be embarrassed if the donor sees her offer being rejected. The likelihood that the proposer's offer is a signal towards the donor is high when the donor is known in the community and has political connections, friends and enemies, as was the case with the Ache (Hill & Gurven 2004).

The same holds for the responder, who can expect that the donor knows what amount he accepted or refused. A responder may also be concerned with creating a reputation of being tough by rejecting a low offer, or seeks social approval by not showing anger or disappointment in public and accepting any offer. Since he can assume he will never find out who the proposer was, and vice versa, the primary target of reputation building appears to be the donor rather than the proposer. In this view, fairness or toughness are signals towards the donor as well as the partner, unlike in two-body analyses of the UG, such as Fehr and Schmidt's (1999) theory of fairness.

The same perspective can be applied to the Dictator Game. The fact that a proposer offers more than zero in the Dictator game has been taken as the demonstration of genuine rather than strategic altruism. Seen as a three-body game, this conclusion does not follow. If the proposer is concerned with her reputation, an anonymous player who cannot identify her consequently cannot promote her reputation, whereas the non-anonymous donor can. This issue seems critical in societies where the donor stands out, in terms of status or knowledge, from the social environment in which the players live.

The three-body view of the UG extends to explanatory attempts in terms of social analogs. Such an explanation was proposed for the Orma, who recognized a similarity between the Public Goods Game and the local contributions Orma households make when the community decides to pursue a common good, such as building a school. For the UG, no such analogy was proposed, and we would be curious to learn from the authors whether they were never observed, or else, what analogies have been made. A three-body view invites looking for analogies with a richer interactive structure: one party donating goods to a second one, while retaining the option to withdraw them if the second party's division of the pie is rejected by a third party. In such cultural analogs, if they exist, the donations could be bribes, gifts, alms, or obligations, or something else. And when the money changes hands from the donor to the proposer, it can change its functional category, such as from a gift to an obligation – but only when considered from the three-body perspective. If people can map the UG into a common analogy, then the variance in the offers (rejections) should decrease, whereas the absolute offers and acceptance levels will still vary with the specific analogy.

How would the behavior be different if the donor provided other goods to the proposer than money? If heuristics for sharing depend on the goods – meat and honey are meticulously shared among the Ache, but goods purchased by money are not (Henrich et al. 2004) – then the observed behavior in the UG should also depend on the kind of pie, not solely on some abstract preferences for selfish or altruistic behavior. In fact, in Lamalera, packs of cigarettes rather than money were used in the UG, and the Lamalera ranked among the top "altruistic" societies. Cigarettes tend to be shared, and this may enhance the appearance of a preference for altruism.

Henrich et al. assume that cultural evolution shapes preferences, yet the alternative to this view is that evolution shapes decision heuristics instead. A tit-for-tat player follows a heuristic, not a preference for altruism or defection, except in the first move. The resulting behavior is based on an interactive strategy, not on preferences that are assumed to be stable like personality traits. We think that the connection between cultural evolution and behavioral economics might be better understood as the shaping of heuristics in the adaptive toolbox. Here, the interpretation of the UG as a three-body transaction provides a new twist to the question of the influence of the environment in which the heuristics of the players are adapted. They may react to the donor as well as to the other player.

What does the Ultimatum Game mean in the real world?

Randolph C. Grace and Simon Kemp

Department of Psychology, University of Canterbury, Christchurch, New Zealand. **randolph.grace@canterbury.ac.nz**
simon.kemp@canterbury.ac.nz

Abstract: The predictive validity of the ultimatum game (UG) for cross-cultural differences in real-world behavior has not yet been established. We discuss results of a recent meta-analysis (Oosterbeek et al 2004), which examined UG behavior across large-scale societies and found that the mean percent offers rejected was positively correlated with social expenditure.

Experiments with ultimatum games (UG) have now been conducted for at least 25 years (cf. Güth et al. 1982). However, there is considerable doubt about which (if any) real world phenomena game behavior might relate to. The predictive validity of the UG as a measure of prosocial behavior is yet to be established.

Henrich et al. are to be congratulated for their efforts in conducting games across an unusually wide range of societies and for clearly showing that rather different average behavior in the game is displayed in these different societies. But which features of the societies are associated with the differences in game behavior? Alternatively, is it possible that differences arise because societies tend to cue different contexts which motivate the respondents to approach the essentially decontextualized UG in different ways? Henrich et al. are unable to answer these questions completely, although they come up with suggestive results and discuss the issues clearly and fairly.

A recent meta-analysis by Oosterbeek et al. (2004) has shown that there are cultural differences in UG behavior. These re-

https://doi.org/10.1017/S0140525X05300148 Published online by Cambridge University Press