Guilt- and Shame-Driven Prosociality Across Societies

Catherine Molho*[1,2], Ivan Soraperra[1], Jonathan F. Schulz[3], and Shaul Shalvi[1]

[1] Center for Research in Experimental Economics and Political Decision Making (CREED), University of Amsterdam, The Netherlands

[2] Department of Experimental and Applied Psychology, VU Amsterdam, The Netherlands

[3] Department of Economics, George Mason University, USA

* Corresponding author: Catherine Molho (c.molho@vu.nl).

Abstract

Impersonal prosociality is considered a cornerstone of a thriving civic society, well-functioning institutions, and a growing economy. Previous research has documented substantial cross-societal variation in prosociality using tasks such as dictator games, where individuals allocate money between themselves and others. In such tasks, individuals typically receive full information about how decisions impact others and make decisions privately. Here, we propose that different societies rely on distinct mechanisms—guilt and internalized norms versus shame and external pressures—to support prosociality. In 20 culturally diverse countries, we will administer dictator games and experimentally induce guilt, by varying information about the consequences of participants' decisions, and shame, by varying observability. Additionally, we will measure guilt- and shame-proneness at the individual and societal level. We will test the hypotheses that activating guilt (by varying information) more strongly increases prosociality among guilt-prone individuals and societies, whereas activating shame (by varying observability) more strongly increases prosociality among shame-prone individuals and societies.

Introduction

Prosociality—a behaviour that benefits others, sometimes at a cost to oneself[1]—is central to well-functioning societies. A large body of work across disciplines, including economics[2–5], psychology[6–10], and anthropology[11,12], has attempted to document and explain variation in prosociality around the world. This work has shown that, across societies, people display different levels of altruistic behavior[2], trust[7], and cooperation[10]. Further, it has revealed substantial cross-societal variation in the strategies that people use to promote cooperation (e.g., via punishment[3,11] or via restricting prosociality to the ingroup[6,8]). Recent theory suggests that cultural evolutionary processes underlie this variation in impersonal prosociality[5,13–15], giving rise to different moral systems—i.e., packages of psychological mechanisms, norms, and institutions that regulate social behaviour.

One important aspect of the moral systems characterizing different societies is the extent to which they rely on guilt versus shame to regulate social interactions. Guilt and shame are both conceptualized as self-conscious moral emotions, yet they have distinct antecedents and consequences[16,17]. Whereas guilt tracks how an individual performs in relation to their own internal moral standards even in private situations[18–20], shame closely tracks social devaluation in the eyes of others[21–25]. Importantly, cultures differ in terms of their predominant emotional response to the commission of violations[13,26–30]—guilt versus shame—and the emphasis they place on internalized moral norms versus external pressures and social reputation, respectively. Recent work has started to uncover the cultural bases of guilt and shame experiences[13,31,32], yet research on how guilt- versus shame-proneness affects prosociality has overwhelmingly focused on individuals from WEIRD societies[33,34]. Here, we propose that societies with different moral systems, i.e., which vary in the emphasis they place on guilt versus shame, rely on distinct mechanisms to support prosociality, and we test this proposition in a culturally diverse set of countries.

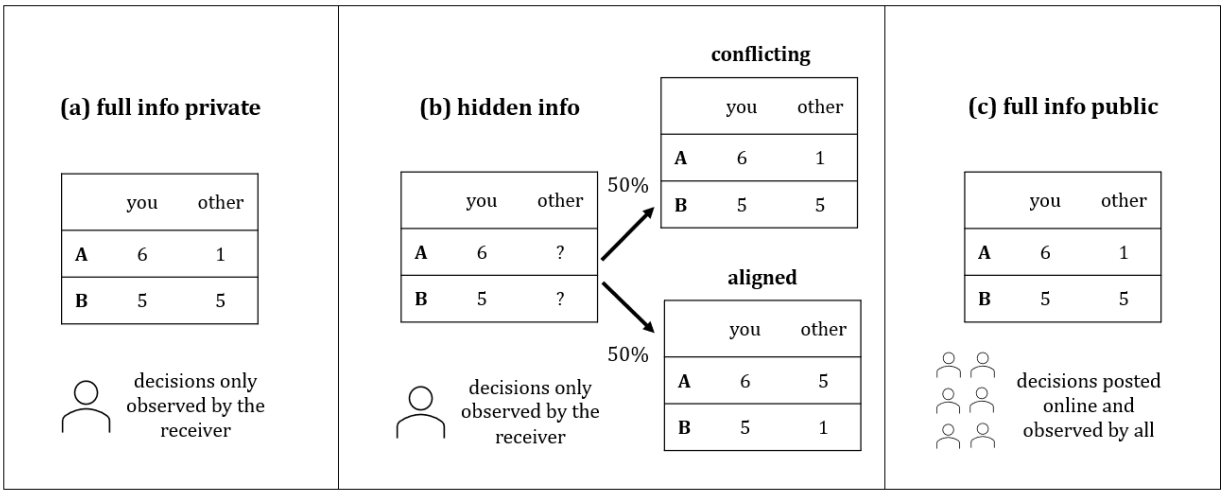
To study cross-societal variation in prosociality, prior research has largely relied on two-person decision-making tasks, such as the prisoner's dilemma, the trust game, and the dictator game. In these tasks, decision-makers typically receive *full information* about the consequences that their own and others' actions have on each person's outcomes, and decisions are made *privately*. To illustrate, in a dictator game, participants are randomly assigned to the roles of allocator and receiver, and the allocators face a decision about how much money to share with the other person. Receivers have no say in this decision. Importantly, allocators know exactly how their decisions will impact receivers, and their decisions are not communicated to anyone outside the decision-making situation. In our study, we build upon the dictator game and introduce two variations that allow us to tease apart the effects of guilt and shame on prosocial decisions (for an Overview, see Figure 1).

First, to activate guilt, we use the 'wilful ignorance' paradigm[35,36], which includes two treatments. The baseline, *full information private* treatment (Figure 1, panel a) consists of a standard binary-choice dictator game. Allocators make a binary decision either in favour of a selfish option (Option A), which yields 6 Monetary Units (MUs) for themselves and 1 MU for the receiver, or a prosocial option (Option B), which yields 5 MUs for both self and other. In this baseline treatment, allocators are informed by default about the negative consequences of choosing Option A for receivers.

Our first treatment variation is the *hidden information* treatment (Figure 1, panel b), where allocators learn with certainty *only* the consequences of each of the two options for their own payoff (i.e., Option A yielding 6 MUs and Option B yielding 5 MUs for the allocator). Receivers' payoffs are clouded in uncertainty. Allocators know that there is a 50% chance that they are in a state with *conflicting interests* and a 50% chance that they are in a state with *aligned interests*. In the *conflicting interests* state, Option A yields 1 MU and Option B yields 5 MUs for the receiver. Thus, in this state, what is best for the allocator (to choose Option A and gain 6 MUs) is worst for the receiver (who gains 1 MU under Option A). In contrast, in the *aligned interests* state, Option A yields 5 MUs

and Option B yields 1 MU for the receiver. Thus, in this state, Option A is best for both allocators (who gain 6 MUs) and receivers (who gain 5 MUs). Allocators have a choice to costlessly obtain information about the state of the world, i.e., the payoff consequences of choosing Options A or B for the receiver, or to proceed without this information.

Figure 1. Overview of experimental treatments. In the baseline, *full information private* treatment (panel a), allocators make decisions in a binary-choice dictator game. Only the allocator and receiver learn about the outcome of the allocator's choice. In the *hidden information* treatment variation (panel b), receivers' payoffs are obscured and allocators can choose to reveal those payoffs or make their decision without knowing them. Again, only the allocator and receiver learn about the outcome of the allocator's choice. In the *full information public* treatment variation (panel c), allocators know receivers' payoffs (as in the baseline). Now, however, the outcome of allocators' choice is publicized on a blog that all participants can see.



Importantly, the treatment variation between *hidden* versus *full information* has consequences for prosocial behaviour. Previous work has shown that a substantial proportion of allocators in the *hidden information* treatment choose to remain ignorant[35], at least partly because ignorance helps reduce the guilt of *knowingly* making a selfish decision[37]. That is, in the baseline, *full information private* treatment, allocators know how their actions affect receivers and have no option to alleviate guilt via ignorance. In contrast, in the *hidden information* treatment they can avoid learning the negative consequences of their actions. As a result, they can make selfish choices while alleviating

guilt[38] and protecting their self-image[36,39]. Indeed, a recent meta-analysis[37] shows a robust prosociality gap between the two treatments, with a larger proportion of allocators making the selfish choice under *hidden information*. To date, however, all studies documenting this prosociality gap have been conducted in a limited number of WEIRD societies (i.e., France, Germany, Netherlands, Norway, Switzerland, and the United States[37]).

Here, we will collect data in a culturally diverse set of 20 countries across all inhabited continents (Africa, North and South America, Asia, Australia, and Europe). Following the rationale above, we will use the comparison between the *full information private* and the *hidden information* treatments as an indication of the extent to which prosociality is guilt-driven (see Figure 2). As described in Table 3, we will then test the hypothesis that:

H_1: The fraction of prosocial choices will be higher in the *full information private* treatment compared to the *hidden information* treatment.

Second, we introduce a *full information public* treatment (see Figure 1, panel c) to activate shame via reputational concerns. As noted earlier, prosocial decisions have typically been studied in private settings, in which allocators' choices are not communicated to anyone outside the decision-making situation, as is the case in our *hidden information* and *full information private* treatments. In our second treatment variation, *full information public,* individuals' decisions become observable. Following prior work[8,40], decisions in this public treatment will be publicized online under a pseudonym and will be accessible by all study participants. Importantly, this experimental manipulation has been shown to increase prosociality[8], partly via increasing individuals' concern for their reputation in the eyes of others. Here, we will use the comparison between the *full information private* and *full information public* treatments to examine the extent to which prosociality is shame-driven (see Figure 2) and, as described in Table 3, test the hypothesis that:

H_2: The fraction of prosocial choices will be higher in the *full information public* treatment compared to the *full information private* treatment.

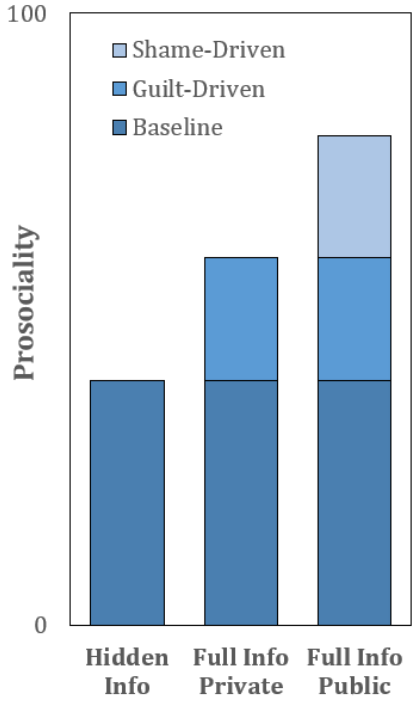


Figure 2. Schematic representation of the hypothesized effects of guilt-related and shame-related factors on the fraction of prosocial choices. The gap between the *hidden information* (Hidden Info) treatment and the *full information private* (Full Info Private) treatment provides an estimate of guilt-driven prosociality. The gap between the *full information private* treatment and the *full information public* (Full Info Public) treatment provides an estimate of shame-driven prosociality.

Activating internalized guilt and external shame should increase prosociality, but the strength of these effects should depend additionally on individual and cultural characteristics. At the individual level, people differ in their proneness to experience guilt and shame in response to committing transgressions[17,19]. And at the societal level, while guilt is the predominant emotional response to committing offenses in some cultures, shame is the predominant emotional experience in others[13,14,26,27,30]. Here, we test how guilt- and shame-activating mechanisms (namely, *information* and *observability*) increase prosociality as a function of individual and cultural differences in guilt- versus shame-proneness. At the individual level, we assess guilt- and shame-proneness using the well-validated GASP scale[32,41] and focus on the difference between guilt- and shame-proneness scores. At the cultural level, we compute averages of guilt- and shame-proneness in each country using the same scale, and again focus on the difference score of the two.

We first examine whether the effect of activating guilt via providing information about the negative consequences of one's choices (*full information private* versus *hidden information*) differs depending on guilt- versus shame-proneness. We propose that prosocial choices among guilt-prone

(versus shame-prone) individuals and cultures will be more strongly based on a desire to fulfil internal moral obligations and maintain a positive self-image (rather than social reputation). If that is the case, then activating guilt by providing information about the negative consequences of selfishness should have a stronger positive effect on prosociality among guilt-prone (versus shame-prone) individuals and cultures. And vice-versa, allowing individuals to alleviate guilt by avoiding information about the negative consequences of their behaviour should have a stronger negative effect on prosociality among guilt-prone (versus shame-prone) individuals and cultures. Based on these ideas and, as described in Table 3, we test the hypothesis that:

H_3: Guilt-prone (versus shame-prone) individuals and cultures will show more guilt-driven prosociality. That is, the gap in prosocial choices between the *hidden information* and the *full information private* treatments will be larger among guilt-prone (versus shame-prone) individuals and cultures.

Next, we examine whether the effect of activating shame via observability (*full information private* versus *full information public*) differs depending on guilt- versus shame-proneness. We propose that prosocial choices among shame-prone (versus guilt-prone) individuals and cultures will be more strongly based on a desire to abide by external moral norms and maintain a positive social reputation. If true, then activating shame by making individuals' decisions publicly observable should have a stronger positive effect on prosociality among shame-prone (versus guilt-prone) individuals and cultures. And vice versa, allowing individuals to escape shame by keeping decisions private should have a stronger negative effect on prosociality among shame-prone (versus guilt-prone) individuals and cultures.

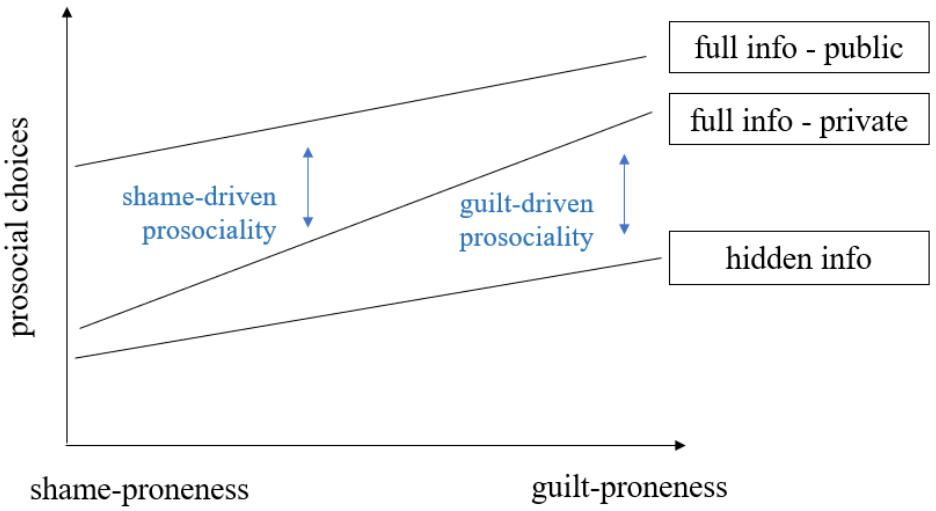
Based on these ideas and, as described in Table 3, we test the hypothesis that:

H_4: Shame-prone (versus guilt-prone) individuals and cultures will show more shame-driven prosociality. That is, the gap in prosociality between the *full information private* and the *full*

information public treatments will be larger among shame-prone (versus guilt-prone) individuals and cultures.

Figure 3 provides a schematic depiction of our hypotheses regarding the effects of guilt-proneness (versus shame-proneness) on prosociality in the *hidden information* and the *full information private* treatments, as well as regarding the effects of guilt-proneness (versus shame-proneness) on prosociality in the *full information private* and the *full information public* treatments.

Figure 3. Schematic depiction of hypothesized effects of guilt-proneness versus shame-proneness on prosociality. Guilt-driven prosociality refers to the gap in prosocial choices between the *hidden information* and *full information private* treatments (H_1), which we expect to increase with guilt- versus shame-proneness at the individual and at the societal level (H_3). Shame-driven prosociality refers to the gap in prosocial choices between the *full information private* and *full information public* treatments (H_2), which we expect to decrease with guilt- versus shame-proneness at the individual and at the societal level (H_4).



Note that our hypotheses (H_3 and H_4) concern how guilt-proneness versus shame-proneness relate with the gap in prosocial choices across treatments. That said, we depict prosociality as increasing with guilt-proneness (versus shame-proneness), based on previous evidence suggesting that guilt-proneness is positively associated with prosociality[33]. At the individual level, experiencing guilt has been associated with more prosocial, conciliatory responses following the commission of a

violation, whereas experiencing shame has been primarily associated with withdrawal responses[16,17] (but see[42,43]). That said, the negative behavioral consequences of shame may not generalize across cultural contexts[44]. At the cultural level, guilt-proneness has been seen as part of a broader culturally evolved moral system—that encompasses universal moral values, belief in moralizing gods, and third-party punishment—underlying impersonal prosociality[13,14].

In addition to testing these a priori formulated hypotheses (see Table 3 for additional details and Table S1 for an overview), we collect additional measures, which can be used to explain heterogeneity in our treatment effects across countries. These measures, including cultural tightness versus looseness[45,46], generalized trust and ingroup/outgroup trust (from the World Values Survey; WVS), and family ties[47] and network embeddedness, have previously been linked to cross-societal variation in impersonal prosociality[5,8,10,14].

In sum, the current work provides a comprehensive cross-cultural investigation of how prosociality is regulated by guilt and internalized norms versus shame and external pressures. To pinpoint the psychological mechanisms underlying when and how people engage in impersonal prosociality, we experimentally vary guilt- and shame-related factors (i.e., *information* and *observability*) and we measure individual and cultural differences in guilt-proneness versus shame-proneness. These measures allow us to disentangle the effects of guilt-proneness versus shame-proneness at the individual level and at the cultural level (i.e., the effect of living in a society where people are on average more or less guilt-prone versus shame-prone). Further, by collecting data in 20 diverse countries across five continents, we contribute to addressing a WEIRD bias in psychological research and ensure representation of understudied populations, as well as heterogeneity in the psychological traits and behavioural phenomena of interest.

Methods

Ethics information

This research complies with ethical regulations for research conducted with human participants. The study protocol has been approved by the ethics review board (reference #EC20220711060717) of the Center for Research in Experimental Economics and Political Decision Making, University of Amsterdam. Informed consent will be obtained from all respondents prior to participation. Participants will receive a baseline compensation for their participation, and they will receive a bonus based on their first decision in the experiment (see Design for details) and one additional randomly selected decision. To ensure comparability across countries, all payoffs will be presented in Monetary Units (MUs), and the conversion between MUs and money will be calculated based on the purchasing power adjusted exchange rate published by the World Bank.

Sampling plan

We plan to recruit participants from 20 countries around the world (see Table 1) via the panel company Toluna. We selected countries so as to ensure that we cover the full range of cross-country variation in impersonal prosociality (i.e., altruism as measured in the Global Preferences Survey[2]). Further, our selection includes countries that show substantial cultural variation, as indicated by well-established measures of cultural differences (individualism versus collectivism[48]; kinship intensity[5], and cultural distance from the USA[49]). Finally, we aimed to ensure a good representation of different geographical regions and different cultural groups (as clustered based on WVS data[50]).

In each country, we aim to recruit a sample size of N = 390 participants, for a total of N = 7,800 participants across the 20 selected countries. We used the MESS library[51] in R[52] to conduct sensitivity power analyses and get an indication of the smallest effect size we would be able to detect when estimating the difference between the proportion of prosocial choices across different treatments. First, we conducted a sensitivity power analysis for a one-tailed test of the within-subjects difference in the proportion of prosocial choices between the *full information private* and the *hidden information* treatments. We used a McNemar test, and set the α level to 0.05 and power to 0.95. Based on the results of the power analysis, a sample size of N = 390 allows us to detect a difference

of about 8 percentage points in the proportion of prosocial choices in each country, which is lower than the treatment effect of *information* (*full information private* versus *hidden information*) on prosocial choices as estimated in a recent meta-analysis (difference in prosocial choices of 15.6 percentage points)[37]. Further details can be found in the Supplementary Information.

Table 1. List of selected countries. The table provides information on the cultural group classification for each country (based on the Inglehart–Welzel Cultural Map – World Values Survey 7), as well as quotas matching our samples to demographic characteristics of the underlying populations.

Country ISO	Country name	Cultural group	Demographic quotas
ARG	Argentina	Latin America	Age, Gender
AUS	Australia	English Speaking	Age, Gender, Region
BRA	Brazil	Latin America	Age, Gender, Region
CHN	China	Confucian	Age, Gender, Region
COL	Colombia	Latin America	Age, Gender
DEU	Germany	Protestant Europe	Age, Gender, Region
EGY	Egypt	African-Islamic	Age, Gender, Region
IDN	Indonesia	African-Islamic	Age, Gender
IND	India	West and South Asia	Age, Gender, Region
ITA	Italy	Catholic Europe	Age, Gender, Region
JPN	Japan	Confucian	Age, Gender, Region
KEN	Kenya	African-Islamic	Age, Gender, Region
KOR	South Korea	Confucian	Age, Gender, Region
MAR	Morocco	African-Islamic	Age, Gender, Region
MEX	Mexico	Latin America	Age, Gender, Region
NGA	Nigeria	African-Islamic	Age, Gender, Region
NLD	Netherlands	Protestant Europe	Age, Gender, Region
PER	Peru	Latin America	Age, Gender
TUR	Turkey	African-Islamic	Age, Gender, Region
USA	United States of America	English-Speaking	Age, Gender, Region

Second, we conducted a sensitivity power analysis for a one-tailed test of the within-subjects difference in the proportion of prosocial choices between the *full information private* and the *full information public* treatments. Again, we used a one-tailed McNemar test of the within-subjects difference in prosocial choices across the two treatments and set the α level to 0.05 and power to 0.95. As above, the sample size of $N = 390$ participants allows us to detect a difference of about 8 percentage points in the proportion of prosocial choices in each country, which is lower than the

treatment effect of *observability* (*full information private* versus *full information public*) as estimated in a recent meta-analysis on the effect of observability on prosociality (r = 0.14 converted into a difference in prosocial choices of 11.5 percentage points)[53].

Within each country, we aim to recruit samples that are matched with the underlying populations in terms of gender, age (18-65 years old), and region (when possible; see Table 1). We will exclude participants who give incomplete responses, as well as participants who spend half of the median completion time or less to complete the study (indicating inattentiveness). There are no other exclusion criteria. The recruitment company will guarantee the pre-specified number of complete responses per country.

Design

Participants will make a series of decisions in binary-choice dictator games across multiple rounds. Importantly, they will receive instructions and make their decision in the first round before learning about the rest of the experiment. This setup allows us to treat the first round as a one-shot interaction, ensuring comparability of our findings to those of other experiments[35,54]. Further, it allows for a between-subjects comparison of decisions across treatments that is not subject to order effects.

In the first round, participants will be randomly assigned to make a decision in one of the three treatments. The first baseline *full information private* treatment (Figure 1, panel a; N = 130) is a standard binary-choice dictator game in which allocators receive full information about the consequences of their actions and make their decisions privately. In the second *hidden information* treatment (Figure 1, panel b; N = 130), allocators do not initially receive information about the consequences of their choice, but can choose to acquire such information by clicking a button. Finally, the third *full information public* treatment (Figure 1, panel c; N = 130) is the same as the baseline (allocators receive full information on the consequences of their actions), though, with the difference that their choice is observable by third parties. All first-round decisions will involve

conflicting interests, and will be made in the payoff set that we outlined earlier (see Figure 1) and which was used by Dana and colleagues[35]. Further, first-round decisions will always be payoff-relevant.

After the first round, participants will make additional decisions in a within-subjects experimental design such that each participant will take part in all three different treatments (for a similar design see[54,55]). As mentioned above, prior to making these additional decisions, participants make a one-shot decision in one of our three treatments of interest. Therefore, participants are exposed to one of three different decision orders. To account for potential order effects, we will check if the first-round treatment type affects subsequent decisions and control for first-round treatment type in our within-subject analyses.

To experimentally manipulate *information* (i.e., a guilt-activating factor), we will ask participants to make decisions in either (a) a *full information private* treatment or (b) a *hidden information* treatment. To experimentally manipulate *observability* (i.e., a shame-activating factor), we will further ask participants to make decisions in (3) the *full information public* treatment. In this treatment, participants will know that their decisions will be publicized on a website under a pseudonym of their choosing right after they complete the study (see[8]). Besides the first-round decision, which will always be payoff-relevant, one additional randomly chosen decision out of all the decisions each participant made will be paid out at the end of the experiment. Decisions will be incentivized *and* consequential, such that other players will be real. We will use the strategy method and ask participants to make decisions knowing that they will be randomly chosen as either allocators or receivers. Data collection and analysis will not be performed blind to the treatments of the experiment.

Across rounds, participants will be faced with decisions in two different types of situations. We follow Dana and colleagues[35] and have participants make decisions either in situations with *conflicting interests*, where the option that maximizes the payoff of the allocator (i.e., Option A

which yields 6 MUs) is costly to the receiver (i.e., yields only 1 MU), or in a state with *aligned interests*, where the option that maximizes the payoff of the allocator *also* maximizes the payoff of the receiver (i.e., yields 5 MUs; see Figure 1 and Table 2). In addition to situations with conflicting interests, we include situations with aligned interests (a) to ensure the credibility of the *hidden information* treatment (where participants can reveal information about two potential underlying situations) and (b) as a check that participants understand the incentive structure. For the main part of our experiment, we will use the same incentive structure as Dana and colleagues[35] (see Figure 1 and Table 2). Participants will make a total of 6 decisions in this incentive structure—that is, one decision in each situation (*conflicting interests*; *aligned interests*) in each treatment (*full information private*; *hidden information*; *full information public*). To check the robustness of our results, we will implement additional decisions in different incentive structures. Across all incentive structures, participants will make a total of 14 decisions in our experiment (see a detailed list in Table S2).

Table 2. List of within-subjects decisions across rounds. Participants make decisions in three treatments (*full information private, hidden information*, and *full information public*) and they encounter decisions in two types of choice situations (conflicting versus aligned interests).

Treatment	Situation	Option	Payoffs	
			Self	Other
Full information private	Conflicting interests	A	6	1
		B	5	5
	Aligned interests	A	6	5
		B	5	1
Hidden information	Conflicting interests	A	6	[1]
		B	5	[5]
	Aligned interests	A	6	[5]
		B	5	[1]
Full information public	Conflicting interests	A	6	1
		B	5	5
	Aligned interests	A	6	5
		B	5	1

Note: In situations that involve conflicting interests, Option A benefits the self at the detriment of the other, whereas Option B benefits the other at the detriment of the self. In situations that involve aligned interests, payoffs for the other person are flipped, such that Option A benefits both the self *and* the other, whereas Option B is detrimental for both. The square brackets indicate that participants are not aware of the other's payoff, but they may choose to reveal it.

Guilt-Driven Prosociality. In each of the countries that we consider (see Sampling Plan), we plan to estimate the absolute gap in prosociality between the *full information private* and *hidden information* treatments. That is, our analyses will focus on the difference between the fraction of prosocial choices in the *full information private* treatment minus the fraction of prosocial choices in the *hidden information* treatment, in each country. Our treatment comparisons will only focus on decisions in situations with conflicting interests. To estimate the fraction of prosocial choices in the *hidden information* treatment, we follow the state-of-the-art in the literature on willful ignorance and focus on the full sample of participants (rather than only those participants who choose to reveal information). This allows for a clean comparison of the fraction of prosocial choices between the *full information private* and the *hidden information* treatments.

Shame-Driven Prosociality. In each of the countries that we consider (see Sampling Plan), we plan to estimate the absolute gap in prosociality between the *full information private* and *full information public* treatments. That is, our analyses will focus on the difference between the fraction of prosocial choices in the *full information public* treatment minus the fraction of prosocial choices in the *full information private* treatment, in each country.

After participants complete all rounds of decisions, they will be asked to complete a series of questionnaires:

Guilt and Shame Proneness. We will measure individual differences in guilt-proneness and shame-proneness using the Guilt and Shame Proneness scale (GASP scale[41]). This is a scenario-based 16-item scale that contains two guilt sub-scales that assess (a) negative behaviour evaluations and (b) repair action tendencies following *private transgressions*, and two shame sub-scales that assess (a) negative self-evaluations and (b) withdrawal action tendencies following *publicly exposed transgressions*. Because our focus here is on the *relative* tendency to experience guilt versus shame across individuals and cultures, we will calculate a difference score of guilt- versus shame-proneness, by subtracting the aggregate of the shame sub-scale measuring self-evaluations from the aggregate

of the guilt sub-scale measuring behaviour evaluations. In supplementary analyses, we will also take into account differences in the tendency to engage in different actions following a transgression, by subtracting the aggregate of the all shame-proneness items (tapping negative self-evaluations and withdrawal action tendencies) from the aggregate of all guilt-proneness items (tapping negative behaviour evaluations and repair action tendencies). All indices of guilt- versus shame-proneness will be demeaned and standardized to range between 0 and 1. To assess measurement invariance of the GASP scale across the societies that we consider, we will test for configural, metric, and scalar invariance following the approach described in Young and colleagues[31].

Perceived Reputational Consequences. We will use three items to assess the extent to which selfish behaviour is considered to have negative reputational consequences across different societies. Specifically, after participants make all of their decisions, we will ask them "How shameful do you think it would be if you did not help another person and your choice was posted publicly on a website?" (1 = not at all; 7 = extremely). Additionally, we will ask two questions about the perceived reputational consequences of selfish behaviour among their family and among other participants: "How bad do you think it would be for your reputation if your family found out that you did not help another person?" and " "How bad do you think it would be for your reputation if other participants found out that you did not help another person?" (1 = not at all; 7 = extremely).

Cultural Tightness/Looseness. We will measure cultural tightness versus looseness—i.e., perceptions of the strength of social norms and tolerance of deviant behaviour—across countries, using six questions developed by Gelfand and colleagues[45,46].

Trust. We will measure generalized trust, as well as trust toward specific groups, using items from the latest wave of the World Values Survey (2017-2021). Specifically, to measure generalized trust, we will ask participants to answer the question: *"Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?"* We will also measure how much people trust their family, neighbourhood, people they know personally, people

they meet for the first time, people of another religion, and people of another nationality to create a measure of in-group versus out-group trust.

Family ties. Following Alesina and Giuliano[47], we will measure the strength of families ties based on three questions from the World Values Survey. Specifically, the first item asks *"For each of the following, indicate how important it is in your life. Would you say it is very important, rather important, not very important, or not at all important?"* People will answer this item with regards to their family. The second item is a question that allows a binary answer: "Which of the following statements best describes your views about parents' responsibilities to their children?" and the answer is either (a) "*Parents have a life of their own and should not be asked to sacrifice their own well-being for the sake of their children*." Or (b) *"Parents' duty is to do their best for their children even at the expense of their own well-being."* The third item is a statement that likewise allows for a binary answer: "With which of these two statements do you tend to agree?" and the answer is either (a) "One does not have the duty to respect and love parents who have not earned it by their behaviour and attitudes." Or (b) "Regardless of what the qualities and faults of one's parents are, one must always love and respect them."

Self-esteem and pride-proneness. We will measure self-esteem using the well-established 10-item Rosenberg Self-Esteem scale[56]. To measure proneness to experiencing pride, we will use the 7-item Authentic Pride scale[57].

Demographics. We will measure demographic information, including gender, age, level of education, and subjective social status (using the ladder method). We will further ask participants about their own and their parents' country of birth, their ethnicity, and their religion, using questions from the World Values Survey. To assess participants' actual, rather than subjectively rated, embeddedness in family networks, we will ask questions concerning household size, the number of children, siblings, and cousins that they have, and whether they share their residence with their parents or guardians. Finally, we will ask about the size of the place they grew up in (up to

2,000 inhabitants; 2,000 to 10,000 inhabitants; 10,000 to 100,000 inhabitants; more than 100,000 inhabitants; more than 500,000 inhabitants) and how often they have moved in the last 10 years.

All materials and survey items will be translated into the language of the participant pool in each country using the back-translation method. For each language, we will aim to have one expert translate the original study materials from English and another expert translate it back to English without access to the original materials. In cases where we are not able to recruit experts for the translations, we will hire professional translators. We will resolve discrepancies in consultation with both translators for each language.

Analysis Plan

Pre-processing steps. Before running analyses, we will exclude any incomplete responses.

Hypothesis 1. To test H_1, that the fraction of prosocial choices will be higher in the *full information private* treatment compared to the *hidden information* treatment, we will focus on a key statistic. Specifically, we will focus on choices made in situations that involve conflicting interests (see Table 2), and calculate the treatment effect as the difference in the proportion of prosocial choices between the *full information private* and *hidden information* treatments in each country. These estimated effects will be aggregated across countries using a meta-analysis. Specifically, we plan to conduct a random-effects meta-analysis to estimate the average level of guilt-driven prosociality, i.e., the average difference in prosocial choices between the *full information private* and *hidden information* treatments. An aggregated effect significantly greater than zero will be interpreted as evidence supporting the hypothesis that prosociality is partially driven by guilt. Furthermore, we will employ a Q test for heterogeneous effects to assess whether guilt-driven prosociality is homogeneous or heterogenous across countries.

Hypothesis 2. To test H_2, that the fraction of prosocial choices will be higher in the *full information public* treatment compared to the *full information private* treatment, we will again focus on choices made in situations that involve conflicting interests (see Table 2), and this time calculate

the treatment effect as the difference in the proportion of prosocial choices between the *full information public* and *full information private* treatments in each country. These estimated effects will be aggregated across countries using a meta-analysis. As above, we plan to conduct a random-effects meta-analysis to estimate the average level of shame-driven prosociality. An aggregated treatment effect significantly greater than zero will be interpreted as evidence supporting the hypothesis that prosociality is partially driven by shame. As above, we will employ a Q test for heterogeneous effects to assess whether shame-driven prosociality is homogeneous or heterogenous across countries.

Note that we test each hypothesis based on the relevant sub-sample of choices. For example, in the case of H_1, we focus only on those choices made in the *hidden information* and *full information private* treatments, while to test H_2, we focus only on those choices made in the *full information public* and *full information private* treatments.

Hypothesis 3. To test $\boldsymbol{H_3}$, that guilt-prone versus shame-prone individuals and cultures will show more guilt-driven prosociality (i.e., a larger gap between prosocial choices made in the *full information private* treatment compared to the *hidden information* treatment), we will run a linear probability regression model, see equation (1), with participants' choice (selfish versus prosocial) as the dependent variable (where $y_{ijk} = 1$ if choice *i* of participant *j* in country *k* is prosocial and $y_{ijk} = 0$ otherwise).

$$(1)\ \boldsymbol{y_{ijk} = \alpha_0 + \alpha_1 \cdot x_{jk} + \alpha_2 \cdot \overline{x_k} + \gamma \cdot dFIPr + \delta_1 \cdot x_{jk} \cdot dFIPr + \delta_2 \cdot \overline{x_k} \cdot dFIPr + \tau_k + \eta_{jk} + \epsilon_{ijk}}$$

As independent variables, this model will include a *full information private* dummy ($dFIPr$ = 0 in the *hidden information* treatment; $dFIPr$ = 1 in the *full information private* treatment), the average guilt-proneness versus shame-proneness at the country level ($\overline{x_k}$), the individual level guilt-proneness versus shame-proneness (x_{jk}), and the interactions between the *full information private* dummy × country-level guilt-proneness versus shame-proneness ($\overline{x_k} \cdot dFIPr$) and *full information*

private dummy × individual-level guilt-proneness versus shame-proneness ($x_{jk} \cdot dFIPr$). Additionally, we will include random effects at the country (τ_k) and at the individual level (η_{jk}). For expected results and interpretation of the different outcomes, see Table 3.

Hypothesis 4. To test ***H₄,*** that shame-prone versus guilt-prone individuals and cultures will show more shame-driven prosociality (i.e., a larger gap between prosocial choices made in the *full information public* treatment compared to the *full information private* treatment), we will run a linear probability regression model (see equation (2) with participants' choice (selfish versus prosocial) as the dependent variable.

(2) $$\boldsymbol{y_{ijk} = \beta_0 + \beta_1 \cdot x_{jk} + \beta_2 \cdot \overline{x_k} + \theta \cdot dFIPu + \varphi_1 \cdot x_{jk} \cdot dFIPr + \varphi_2 \cdot \overline{x_k} \cdot dFIPu + \tau_k + \eta_{jk} + \epsilon_{ijk}}$$

As independent variables, we will include a *full information public* dummy ($dFIPu$ = 0 in the *full information private* treatment; $dFIPu$ = 1 in the *full information public* treatment), the average guilt-proneness versus shame-proneness at the country level ($\overline{x_k}$), the individual level guilt-proneness versus shame-proneness (x_{jk}), and the interactions between the *full information public* dummy × country-level guilt-proneness versus shame-proneness ($\overline{x_k} \cdot dFIPu$) and *full information public* dummy × individual-level guilt-proneness versus shame-proneness ($x_{jk} \cdot dFIPu$). As before, we will include random effects at the country (τ_k) and at the individual level (η_{jk}). For expected results and interpretation of the different outcomes, see Table 3 below.

Note than in the analyses for hypotheses 3 and 4 we will use linear probability models with robust standard errors because interactions in non-linear models have no straightforward interpretation (see[58,59]). In non-linear models, interaction effects are different from the marginal effects of the interaction and their magnitude and sign depend on the other covariates. In the supplementary materials, we will report robustness check of the results using logit and probit models.

Table 3. Design Table.

Question	Hypothesis	Sampling plan (e.g. power analysis)	Analysis Plan	Interpretation given to different outcomes
1. How does activating guilt and self-image concerns (by providing information versus allowing ignorance) influence prosociality across societies?	***H_1:*** Across the countries considered, the fraction of prosocial choices will be higher in the *full information private* treatment compared to the *hidden information* treatment. That is, providing information about the negative consequences of one's actions for others will have a positive effect on prosocial behaviour.	Based on sensitivity power analysis, a sample of N = 390 gives us 95% power to detect an effect as small as RD = 0.08 for a one-tailed test of the difference between the proportion of prosocial choices in the *full information private* and *hidden information* treatments (when the meta-analytic estimate for this treatment difference is RD = 0.156).	Our main focus is on the risk difference in prosocial choices between the *full information private* and *hidden information* treatments. We will conduct a random-effects meta-analysis to estimate the average level of guilt-driven prosociality (i.e., difference in the proportion of prosocial choices between the two treatments above) across countries, and we will obtain an estimate of guilt-driven prosociality in each country as well.	We interpret a statistically significant treatment effect in the predicted direction (see ***H_1***) as evidence that prosociality is (at least partly) driven by guilt and self-image concerns. If we observe heterogeneity in the treatment effect across countries, we will interpret this as evidence of cross-societal variation in the extent to which prosociality is driven by guilt and self-image. If we observe no statistically significant treatment effect, we will test whether the treatment effect is small enough to be considered negligible, using equivalence testing[61]. Looking at the average level of guilt-driven prosociality across countries, we will use the two one-sided tests (TOST) procedure applied to meta-analysis[62]. Since random-effects meta-analyses assume heterogeneous effects across countries, equivalence testing is not clearly extended to this type of analysis. For this reason, we plan to conduct the equivalence test using a fixed-effects meta-analysis. We consider a difference of 5 percentage points as the smallest effect size of interest. If the TOST rejects the null hypothesis of an effect equal to the smallest effect of interest, we will consider the effect to be negligible.

2. How does activating shame and social reputation concerns (by making decisions observable versus keeping them private) influence prosociality across societies?	***H₂:*** Across the countries considered, the fraction of prosocial choices will be higher in the *full information public* treatment compared to the *full information private* treatment. That is, making one's decisions observable will have a positive effect on prosocial behaviour.	Based on a sensitivity power analysis, a sample of N = 390 gives us 95% power to detect an effect as small as *RD* = 0.08 of observability on the fraction of prosocial choices in each country (when the meta-analytic estimate for this treatment difference is equivalent to *RD* = 0.115).	Our focus is on the risk difference in prosocial choices between the *full information private* and *full information public* treatments. We will conduct a random-effects meta-analysis to estimate the average level of shame-driven prosociality (i.e., difference in the proportion of prosocial choices between the two treatments above) across countries, and we will obtain an estimate of shame-driven prosociality in each country as well.	We interpret a statistically significant treatment effect in the predicted direction (see ***H₂***) as evidence that prosociality is (at least partly) driven by shame and social reputation concerns. If we observe heterogeneity in the treatment effect across countries, we will interpret this as evidence of cross-societal variation in the extent to which prosociality is driven by shame and social reputation concerns. If we observe no statistically significant treatment effect, we will test whether the treatment effect is small enough to be considered negligible, using equivalence testing[61]. For the reason explained above, we will use the two one-sided tests (TOST) procedure applied to a fixed-effect meta-analysis[62]. We consider a difference of 5 percentage points as the smallest effect size of interest. If the TOST rejects the null hypothesis of an effect equal to the smallest effect of interest, we will consider the effect to be negligible.
3. Can individual and cultural differences in guilt-proneness versus shame-proneness explain heterogeneity in the prosociality	***H₃:*** Guilt-prone (versus shame-prone) individuals and cultures will show more guilt-driven prosociality. That is, the gap in prosocial choices between the *hidden information* and the *full information private* treatments will be increasing with more guilt-prone (versus	This analysis will make use of the choices in the conflicting interests situations of the *full information private* and the *hidden information* treatments (*N* = 390 x 2 x 20).	We plan to examine how guilt-proneness versus shame-proneness (i.e., the difference score of guilt- and shame-proneness) at the individual and at the country level relate with guilt-driven prosociality. We will run a linear probability regression model (in equation (1), which has participants' choice (selfish versus prosocial) as a dependent variable. As independent variables, the	Our interpretation will focus on the parameters of the *full information private* dummy (γ) and of the interaction terms between the *full information private* dummy and individual- and country-level guilt-proneness versus shame-proneness, δ_1 and δ_2 respectively. We will interpret a non-negative treatment dummy ($\gamma \geq 0$) and a statistically significant positive interaction term of the *full information private*

gap between the *full information private* and *hidden information* treatments (i.e., guilt-driven prosociality)?	shame-prone) individuals and cultures.		model will include a *full information private* dummy, the guilt-proneness versus shame-proneness at the individual and at the country level, as well as the interactions of the *full information private* dummy × country-level guilt-proneness versus shame-proneness and the *full information private* dummy × individual guilt-proneness versus shame-proneness.	dummy with the individual guilt-shame proneness ($\delta_1 > 0$) as evidence supporting our hypothesis that individuals with higher relative levels of guilt-proneness versus shame-proneness show a larger gap in prosocial choices between the *hidden information* and *full information private* treatments. If the interaction term of the *full information private* dummy × individual guilt-proneness versus shame-proneness is not statistically significant, we will conclude that we cannot reject the null hypothesis. We will proceed to test whether the effect of the standardized index of individual guilt-proneness versus shame-proneness on guilt-driven prosociality (i.e., the gap in the proportion of prosocial choices between the *full information private* and *hidden information* treatments) is small enough to be considered negligible, using equivalence testing[61] (with the R library PARAMETERS). We will apply the TOST procedure on the regression coefficient δ_1 considering a smallest effect of interest of 0.05 (which is equivalent to an increase of 5 percentage points when moving from the lowest to the highest level of the guilt-proneness versus shame-proneness index). In addition, we will interpret a statistically significant positive interaction term of the *full information private* dummy × country-level guilt-proneness versus shame-proneness ($\delta_2 > 0$) as evidence that,

				ceteris paribus, living in a country with higher levels of guilt-proneness versus shame-proneness is associated with a larger gap between the *hidden information* and *full information private* treatments. In other words, individuals with a similar level of guilt- versus shame-proneness will show larger differences in prosocial choices in countries relatively more prone to guilt than to shame. If the interaction term of the *full information private* dummy × country-level guilt-proneness versus shame-proneness is not statistically significant, we will conclude that we cannot reject the null hypothesis. We will then proceed to estimate whether the effect of the standardized index of country-level guilt-proneness versus shame-proneness on guilt-driven prosociality (i.e., the gap in prosocial choices between the *full information private* and *hidden information* treatments) is small enough to be considered negligible, using equivalence testing[61] (with the R library PARAMETERS). We will apply the TOST procedure on the regression coefficient δ_2 considering a smallest effect of interest of 0.05 (which is equivalent to an increase of 5 percentage points when moving from the lowest to the highest level of the guilt-proneness versus shame-proneness index).

4. Can individual and cultural differences in guilt-proneness versus shame-proneness explain heterogeneity in the prosociality gap between the *full information private* and *full information public* treatments (i.e., shame-driven prosociality)?	***H₄:*** Shame-prone (versus guilt-prone) individuals and cultures will show more shame-driven prosociality. That is, the gap in the proportion of prosocial choices between the *full information private* and the *full information public* treatments will be larger among relatively more shame-prone (versus guilt-prone) individuals and cultures.	This analysis will make use of the choices in the conflicting interests situations of the *full information private* and of the *full information public* treatments (N = 390 x 2 x 20).	We plan to examine how guilt-proneness versus shame-proneness (i.e., the difference score of guilt- and shame-proneness) at the individual and at the country level relate with shame-driven prosociality. We will run a linear probability regression model (in equation (2)), which has participants' choice (selfish versus prosocial) as a dependent variable. As independent variables, the model will include a *full information public* dummy, the guilt-proneness versus shame-proneness at the individual and at the country level, as well as the interactions of the *full information public* dummy × country-level guilt-proneness versus shame-proneness and the *full information public* dummy × individual guilt-proneness versus shame-proneness.	Our interpretation will focus on the parameters of the *full information public* dummy (θ) and of the interaction terms between the *full information public* dummy and individual- and country-level guilt-proneness versus shame-proneness, φ_1 and φ_2 respectively. We will interpret a non-negative treatment dummy ($\theta > 0$) and a statistically significant negative interaction term of the *full information public* dummy with the individual guilt-proneness versus shame- proneness ($\varphi_1 < 0$) as evidence supporting our hypothesis that individuals with lower relative levels of guilt- versus shame-proneness show a larger gap in prosocial choices between the *full information public* and *full information private* treatments. If the interaction term of the *full information public* dummy × individual guilt-proneness versus shame-proneness is not statistically significant, we will conclude that we cannot reject the null hypothesis. We will proceed to test whether the effect of the standardized index of individual guilt-proneness versus shame-proneness on shame-driven prosociality (i.e., the gap in prosocial choices between the *full information public* and *full information private* treatments) is small enough to be considered negligible, using equivalence testing[61] (with the R library PARAMETERS). We will apply the TOST procedure on the

				regression coefficient φ_1 considering a smallest effect of interest of 0.05 (which is equivalent to a decrease of 5 percentage points in the gap of prosocial choices when moving from the lowest to the highest level of the guilt-proneness versus shame-proneness index). In addition, we will interpret a statistically significant negative interaction term of the *full information public* dummy × country-level guilt-proneness versus shame-proneness ($\varphi_2 < 0$) as evidence that, ceteris paribus, living in a country with higher levels of shame- versus guilt-proneness is associated with a larger gap between the *full information public* and *full information private* treatments. In other words, individuals with a similar level of guilt-proneness versus shame-proneness will show larger differences in prosocial choices in countries relatively more prone to shame than to guilt. If the interaction term of the *full information public* dummy × country-level guilt-proneness versus shame-proneness is not statistically significant, we will conclude that we cannot reject the null hypothesis.. We will then proceed to estimate whether the effect of the standardized index of country-level guilt-proneness versus shame-proneness on shame-driven prosociality (i.e., the gap in prosocial choices between the *full information public* and *full information private* treatments) is small enough to be considered negligible, using equivalence

				testing[61] (with the R library PARAMETERS). We will apply the TOST procedure on the regression coefficient φ_2 considering a smallest effect of interest of 0.05 (which is equivalent to a decrease of 5 percentage points in the gap of prosocial choices when moving from the lowest to the highest level of the guilt-proneness versus shame-proneness index).

References

1. Pfattheicher, S., Nielsen, Y. A. & Thielmann, I. Prosocial behavior and altruism: A review of concepts and definitions. *Current Opinion in Psychology* **44**, 124–129 (2021).
2. Falk, A. *et al.* Global Evidence on Economic Preferences. *The Quarterly Journal of Economics* **133**, 1645–1692 (2018).
3. Herrmann, B., Thöni, C. & Gächter, S. Antisocial Punishment Across Societies. *Science* **319**, 1362–1367 (2008).
4. Gächter, S., Herrmann, B. & Thöni, C. Culture and cooperation. *Phil. Trans. R. Soc. B* **365**, 2651–2661 (2010).
5. Schulz, J. F., Bahrami-Rad, D., Beauchamp, J. P. & Henrich, J. The Church, intensive kinship, and global psychological variation. *Science* **366**, eaau5141 (2019).
6. Dorrough, A. R. & Glöckner, A. Multinational investigation of cross-societal cooperation. *Proc Natl Acad Sci USA* **113**, 10836–10841 (2016).
7. Romano, A., Balliet, D., Yamagishi, T. & Liu, J. H. Parochial trust and cooperation across 17 societies. *PNAS* **114**, 12702–12707 (2017).
8. Romano, A., Sutter, M., Liu, J. H., Yamagishi, T. & Balliet, D. National parochialism is ubiquitous across 42 nations around the world. *Nat Commun* **12**, 4456 (2021).
9. Balliet, D. & Van Lange, P. A. M. Trust, Punishment, and Cooperation Across 18 Societies: A Meta-Analysis. *Perspect Psychol Sci* **8**, 363–379 (2013).
10. Spadaro, G. *et al.* Cross-cultural variation in cooperation: A meta-analysis. *Journal of Personality and Social Psychology* **123**, 1024–1088 (2022).
11. Henrich, J. *et al.* Costly Punishment Across Human Societies. *Science* **312**, 1767–1770 (2006).
12. Henrich, J. *et al.* Markets, Religion, Community Size, and the Evolution of Fairness and Punishment. *Science* **327**, 1480–1484 (2010).
13. Enke, B. Kinship, Cooperation, and the Evolution of Moral Systems*. *The Quarterly Journal of Economics* **134**, 953–1019 (2019).

14. Henrich, J. *The Weirdest People in the World: How the West Became Psychologically Peculiar and Particularly Prosperous*. (Penguin UK, 2020).
15. Henrich, J. & Muthukrishna, M. The Origins and Psychology of Human Cooperation. *Annu. Rev. Psychol.* **72**, 207–240 (2021).
16. Haidt, J. The moral emotions. in *Handbook of affective sciences* 852–870 (Oxford University Press, 2003).
17. Tangney, J. P. & Dearing, R. L. *Shame and Guilt*. (Guilford Press, 2003).
18. Cohen, T. R., Panter, A. T. & Turan, N. Guilt Proneness and Moral Character. *Curr Dir Psychol Sci* **21**, 355–359 (2012).
19. Wolf, S. T., Cohen, T. R., Panter, A. T. & Insko, C. A. Shame Proneness and Guilt Proneness: Toward the Further Understanding of Reactions to Public and Private Transgressions. *Self and Identity* **9**, 337–362 (2010).
20. Bierbrauer, G. Reactions to Violation of Normative Standards: A Cross-Cultural Analysis of Shame and Guilt. *International Journal of Psychology* **27**, 181–193 (1992).
21. Smith, R. H., Webster, J. M., Parrott, W. G. & Eyre, H. L. The role of public exposure in moral and nonmoral shame and guilt. *Journal of Personality and Social Psychology* **83**, 138–159 (2002).
22. Sznycer, D. *et al.* Shame closely tracks the threat of devaluation by others, even across cultures. *PNAS* **113**, 2625–2630 (2016).
23. Sznycer, D. *et al.* Cross-cultural invariances in the architecture of shame. *PNAS* **115**, 9702–9707 (2018).
24. Fessler, D. Shame in Two Cultures: Implications for Evolutionary Approaches. *Journal of Cognition and Culture* **4**, 207–262 (2004).
25. Gao, J., Wang, A. & Qian, M. Differentiating shame and guilt from a relational perspective: a cross-cultural study. *Social Behavior and Personality: An International Journal* **38**, 1401–1408 (2010).

26. Scherer, K. R. & Wallbott, H. G. Evidence for universality and cultural variation of differential emotion response patterning. *J Pers Soc Psychol* **66**, 310–328 (1994).
27. Leung, A. K.-Y. & Cohen, D. Within- and between-culture variation: Individual differences and the cultural logics of honor, face, and dignity cultures. *Journal of Personality and Social Psychology* **100**, 507–526 (2011).
28. Mesquita, B., Boiger, M. & De Leersnyder, J. Doing emotions: The role of culture in everyday emotions. *European Review of Social Psychology* **28**, 95–133 (2017).
29. Boiger, M., Mesquita, B., Uchida, Y. & Feldman Barrett, L. Condoned or Condemned: The Situational Affordance of Anger and Shame in the United States and Japan. *Pers Soc Psychol Bull* **39**, 540–553 (2013).
30. Rodriguez Mosquera, P. M., Manstead, A. S. R. & Fischer, A. H. The role of honour concerns in emotional reactions to offences. *Cognition & Emotion* **16**, 143–163 (2002).
31. de Groot, M. *et al.* Group-based shame, guilt, and regret across cultures. *European Journal of Social Psychology* **51**, 1198–1212 (2021).
32. Young, I. F. *et al.* A multidimensional approach to the relationship between individualism-collectivism and guilt and shame. *Emotion* **21**, 108–122 (2021).
33. Tignor, S. M. & Colvin, C. R. The Interpersonal Adaptiveness of Dispositional Guilt and Shame: A Meta-Analytic Investigation. *Journal of Personality* **85**, 341–363 (2017).
34. Thielmann, I., Spadaro, G. & Balliet, D. Personality and prosocial behavior: A theoretical framework and meta-analysis. *Psychological Bulletin* **146**, 30–90 (2020).
35. Dana, J., Weber, R. A. & Kuang, J. X. Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory* **33**, 67–80 (2007).
36. Grossman, Z. & van der Weele, J. J. Self-Image and Willful Ignorance in Social Decisions. *Journal of the European Economic Association* **15**, 173–217 (2017).
37. Vu, L., Soraperra, I., Leib, M., van der Weele, J. J. & Shalvi, S. Willful Ignorance: A Meta-Analysis. *Working paper* (2022).

38. Hertwig, R. & Engel, C. Homo Ignorans: Deliberately Choosing Not to Know. *Perspect Psychol Sci* **11**, 359–372 (2016).
39. Bénabou, R. & Tirole, J. Incentives and Prosocial Behavior. *American Economic Review* **96**, 1652–1678 (2006).
40. Milinski, M., Semmann, D., Krambeck, H.-J. & Marotzke, J. Stabilizing the Earth's climate is not a losing game: Supporting evidence from public goods experiments. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 3994–3998 (2006).
41. Cohen, T. R., Wolf, S. T., Panter, A. T. & Insko, C. A. Introducing the GASP scale: A new measure of guilt and shame proneness. *Journal of Personality and Social Psychology* **100**, 947–966 (2011).
42. de Hooge, I. E., Breugelmans, S. M. & Zeelenberg, M. Not so ugly after all: When shame acts as a commitment device. *Journal of Personality and Social Psychology* **95**, 933–943 (2008).
43. Declerck, C. H., Boone, C. & Kiyonari, T. No place to hide: When shame causes proselfs to cooperate. *The Journal of Social Psychology* **154**, 74–88 (2014).
44. Wong, Y. & Tsai, J. Cultural models of shame and guilt. in *The self-conscious emotions: Theory and research* 209–223 (Guilford Press, 2007).
45. Gelfand, M. J. *et al.* Differences Between Tight and Loose Cultures: A 33-Nation Study. *Science* **332**, 1100–1104 (2011).
46. Harrington, J. R. & Gelfand, M. J. Tightness–looseness across the 50 united states. *PNAS* **111**, 7990–7995 (2014).
47. Alesina, A. & Giuliano, P. The power of the family. *J Econ Growth* **15**, 93–125 (2010).
48. Hofstede, G. *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*. (SAGE Publications, 2001).
49. Muthukrishna, M. *et al.* Beyond Western, Educated, Industrial, Rich, and Democratic (WEIRD) Psychology: Measuring and Mapping Scales of Cultural and Psychological Distance. *Psychol Sci* **31**, 678–701 (2020).

50. Inglehart, R. & Baker, W. E. Modernization, Cultural Change, and the Persistence of Traditional Values. *American Sociological Review* **65**, 19–51 (2000).
51. Claus Thorn Ekstrøm. MESS: Miscellaneous Esoteric Statistical Scripts. (2022).
52. R Core Team. R: A language and environment for statistical computing. (2022).
53. Bradley, A., Lawrence, C. & Ferguson, E. Does observability affect prosociality? *Proc. R. Soc. B.* **285**, 20180116 (2018).
54. Momsen, K. & Ohndorf, M. When do people exploit moral wiggle room? An experimental analysis of information avoidance in a market setup. *Ecological Economics* **169**, 106479 (2020).
55. Feiler, L. Testing models of information avoidance with binary choice dictator games. *Journal of Economic Psychology* **45**, 253–267 (2014).
56. Rosenberg, M. *Society and the Adolescent Self-Image*. (Princeton University Press, 2015).
57. Tracy, J. L. & Robins, R. W. The psychological structure of pride: A tale of two facets. *Journal of Personality and Social Psychology* **92**, 506–525 (2007).
58. Gomila, R. Logistic or linear? Estimating causal effects of experimental treatments on binary outcomes using regression analysis. *Journal of Experimental Psychology: General* **150**, 700–709 (2021).
59. Ai, C. & Norton, E. C. Interaction terms in logit and probit models. *Economics Letters* **80**, 123–129 (2003).
60. van der Weele, J. J., Soraperra, I., Villeval, M. C. & Shalvi, S. The Social Construction of Ignorance: Experimental Evidence. *Working paper*, hal-03725590 (2022).
61. Lakens, D., Scheel, A. M., & Isager, P. M. Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science* **1**, 259-269 (2018).
62. Lakens, D.. Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. *Social Psychological and Personality Science* **8**, 355-362 (2017).

Data availability

We commit to sharing the raw data and full materials for our study upon acceptance of the Stage 2 manuscript. In accordance with the data management protocol in our institution, we will make data and materials available on the institutional open access repository on FigShare.

Code availability

We commit to sharing all code upon acceptance for publication of the Stage 2 manuscript. Code will be uploaded to our institutional open access repository on FigShare.

Acknowledgements

This research is financially supported by the European Research Council (ERC-CoG-865931) and the Dutch Research Council (NWO; Vi.Vidi.195.137) grants awarded to Shaul Shalvi. The funders have/had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Competing interests

The authors declare no competing interests.

Supplementary information

The table below presents an overview of our hypotheses and study goals.

Table S1. Overview of hypotheses.

#	Hypothesis	Goal
H_1	Providing information (*full information private* versus *hidden information*) about the negative consequences of allocators' actions for others will increase prosocial choices.	Testing the information effect[35,37] in diverse societies
H_2	Making allocators' actions observable (*full information public* versus *full information private*) by third parties will increase prosocial choices.	Testing the observability effect[53] in diverse societies
H_3	The effect of providing information (*full information private* versus *hidden information*) on prosociality will be stronger among guilt-prone individuals and cultures.	Explaining heterogeneity in information effect
H_4	The effect of observability (*full information public* versus *full information private*) on prosociality will be stronger among shame-prone individuals and cultures.	Explaining heterogeneity in observability effect

Robustness checks. To check the robustness of our findings beyond the specific payoff set used by Dana et al.[35], we will ask participants to make decisions in two additional payoff sets (for an overview, see Table S2). Data from these additional choice situations will be analysed separately.

Table S2. List of within-subjects decisions across rounds. Participants make decisions in three treatments (*full information private, hidden information*, and *full information public*). Across rounds, they face decisions in different choice situations and payoff sets. Payoff set 1 is the same as in Dana and colleagues[35]. Shaded areas highlight how payoff sets 2 and 3 differ from payoff set 1.

Treatment	Payoff set	Situation	Option	Payoffs	
				self	other
Full information private	**1**	Conflicting interests	A	6	1
			B	5	5
		Aligned interests	A	6	5
			B	5	1
	2	Conflicting interests	A	9	1
			B	5	5
	3	Conflicting interests	A	6	4
			B	5	5
Hidden information	**1**	Conflicting interests	A	6	[1]
			B	5	[5]
		Aligned interests	A	6	[5]
			B	5	[1]
	2	Conflicting interests	A	9	[1]
			B	5	[5]
		Aligned interests	A	9	[5]
			B	5	[1]
	3	Conflicting interests	A	6	[4]
			B	5	[5]
		Aligned interests	A	6	[5]
			B	5	[4]
Full information public	**1**	Conflicting interests	A	6	1
			B	5	5
		Aligned interests	A	6	5
			B	5	1
	2	Conflicting interests	A	9	1
			B	5	5
	3	Conflicting interests	A	6	4
			B	5	5

Note: In situations that involve conflicting interests, Option A benefits the self at the detriment of the other, whereas Option B benefits the other at the detriment of the self. In situations that involve aligned interests, payoffs for the other person are flipped, such that Option A benefits both the self *and* the other, whereas Option B is detrimental for both. The square brackets indicate that participants are not aware of the other's payoff, but they may choose to reveal it.

Power analyses. We conducted a sensitivity power analysis based on a one-tailed McNemar test to test differences in the proportion of prosocial choices ($RD = \pi_{1\cdot} - \pi_{\cdot 1} = \pi_{12} - \pi_{21}$ in Table S3). R code to replicate this analysis is available on the Open Science Framework: https://osf.io/tdcpn/?view_only=005eb5c9d5d64f80a05fe0e376878246

Table S3. Contingency table of prosocial choices in comparison treatment and baseline treatment.

	Prosocial choice in comparison treatment	Selfish choice in comparison treatment	
Prosocial choice in baseline	π_{11}	π_{12}	$\pi_{1\cdot}$
Selfish choice in baseline	π_{21}	π_{22}	$\pi_{2\cdot}$
	$\pi_{\cdot 1}$	$\pi_{\cdot 2}$	

Assuming an α level of 0.05 and 95% power, with a sample of 390 participants per country we can detect a risk difference of at least 0.073 when the proportion of discordant pairs $\pi_{12} + \pi_{21}$ is set to 0.191, an estimated value based on the within-subject data in[60]. Figure S1 shows the minimum risk difference that is detectable for different statistical power levels as a function of the proportion of discordant pairs. The figure demonstrates that we obtain 95% power to detect effect sizes that are smaller than current meta-analytic estimates of the effects of *information* (RD = 0.156)[37] and *observability* (RD = 0.115)[53] on prosocial choices, for a reasonable range of values of the fraction of discordant pairs (0.20 to 0.30).

Figure S1. Sensitivity power analysis for a one-tailed McNemar test of within-subject differences in the proportion of prosocial choices. The analysis shows the minimum detectable effect size as a function of the proportion of discordant pairs with $N = 390$, $\alpha = 0.05$, and different levels of power.