The Causal Effect of Cultural Identity on Cooperation

Jeffrey V. Butler and Dietmar Fehr*

February 17, 2023

Abstract

The impact of culture on non-kin cooperation has been singled out as critical for economic activity. However, causal evidence of culture's influence on cooperation remains scant. In this paper we provide such evidence, focusing on two key components of culture: preferences and beliefs. Adopting the view that culture is one aspect of an individual's multi-faceted self-concept (identity) we conduct an experiment with foreign- and US-born Chinese immigrants at a large US public university. In a two-by-two design, we exogenously vary: i) the salience of participants' American or Chinese cultural identities; and ii) the capacity for culture to affect beliefs by randomly providing previous-session cooperation-rate information. Comparing behavior across cultures and information conditions, our results suggest a prominent role for both preferences and beliefs. In particular, we find that culture's effects through beliefs are as important as its effects through preferences.

Keywords: cooperation, culture, identity, beliefs, preference, experiment
JEL classification numbers: *C91, D01, O10, P16, Z10*

*Butler: University of California, Merced, jbutler6@ucmerced.edu. Fehr: Heidelberg University, Alfred-Weber-Institute for Economics, dietmar.fehr@awi.uni-heidelberg.de. We thank George A. Akerlof, Dan Benjamin, Johannes Buggle, Yan Chen, Alain Cohn, Paola Giuliano, Aidin Hajikhameneh, Erik Kimbrough, Noam Yuchtman and Xiaoyu Zhou, and participants at the ESA Meetings in Heidelberg and Tucson, the Arne Ryde Workshop in Lund, the SPI Conference in Chicago, BABEEW at the University of San Francisco, the ERINN Workshop at Oxford University, the NCBEE in Kiel, Florida State University, Georgia State University, and the University of California, Merced for helpful comments that have greatly improved the manuscript. Yannick Reichlin and Lennard Naumann provided excellent research assistance. We are grateful to the Xlab Berkeley team for helping us conduct the experiments. Dietmar Fehr gratefully acknowledges financial support from the Berlin Social Science Center (WZB). This study was approved by the Institutional Review Board (IRB) at UC Berkeley (2014-03-6094). The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

1 Introduction

There has long been a perception that culture plays a crucial role in a variety of economic outcomes. The emerging consensus in economics on a concrete definition of culture as beliefs and values shared among people groups has led to rapid growth in the economics literature on culture (Guiso, Sapienza and Zingales, 2006; Bisin and Verdier, 2008; Tabellini, 2010; Fernández, 2011).[1] While the question of *whether* culture matters for economic outcomes is seemingly settled (Fernández and Fogli, 2006; Tabellini, 2010; Alesina, Giuliano and Nunn, 2013; Algan and Cahuc, 2013; Lowes et al., 2017), the channels through which culture affects behavior and economic outcomes is an important open question.

In this paper, we take a step forward in this question and study experimentally the two channels suggested by the definition above: beliefs and values. In particular, we provide evidence on their relative importance in culture's causal impact on non-kin cooperation. We focus on stranger cooperation because it is a question of perennial interest across multiple disciplines which economists have argued facilitates the well functioning of organizations, economies and societies (Knack and Keefer, 1997; Algan and Cahuc, 2013; Alesina and Giuliano, 2015; Enke, 2019). Conceptually, we conjecture that culture may affect behavior directly through preferences by, for example, instilling cooperation as a virtue (Tabellini, 2008), or indirectly, by coloring beliefs about others' preferences and behavior. Beliefs may be colored in various ways. For instance, culture may transmit beliefs about how particular groups or individuals will behave.[2] Another way culture may color beliefs is by influencing which contextual cues receive attention, affecting how unfamiliar others are categorized. By affecting categorization, culture may, in turn, affect beliefs about those others.

As an example of the latter, there are countless situations featuring some contextual cues that, if attended to, could lead individuals to perceive strangers as similar to themselves (human, common fate) and some cues that could lead to a perception of strangers as being different (gender, ethnicity, role). Research in cultural psychology suggests that Easterners perceive relationships more readily than Westerners and, consequently, tend to categorize objects based on relationships (monkey, banana) rather than properties (monkey, panda) (Ji, Zhang and Nisbett, 2004). Contextual cues related to relationships may therefore receive more attention from Easterners than from Westerners. Similarly, social relationships, with a particular focus on delineating the in-group, occupy a central role in collectivist cultures but not in individualist cultures, which in contrast prioritize self over social groups (Hwang, 1987; Triandis, 1995).[3] Moreover, across cultures, once categorization occurs intergroup differences tend to be subconsciously exaggerated (Haslam and Ellemers, 2005; Butler, 2018), perhaps amplifying these indirect beliefs channels.

[1] For example, Guiso, Sapienza and Zingales (2006, p. 23) define culture as "those customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation." Similarly, Bisin and Verdier (2008) refer to culture as "preferences, beliefs, and norms that govern human behavior" while Fernández (2011, p.482) provides a working definition of culture as a "distribution of social preferences and beliefs." Tabellini (2010) measures culture by aggregating "specific indicators of values and beliefs."

[2] The literature on trust may be illustrative, as parents and other authorities often teach children about "stranger danger," which can be thought of as culturally transmitted (pessimistic) beliefs about the intentions, traits and behavior of a particular category of people, strangers (e.g., Binzel and Fehr, 2013; Butler et al., 2016; Butler, Giuliano and Guiso, 2015, 2016).

[3] For example, Triandis (1995, p. 6) asserts that, in contrast to individualism, collectivism entails an "emphasis on the views, needs and goals of the in-group rather than the self," and that collectivism in China can be traced back to the "extreme cooperation and coordination" needed to produce public projects that protected Chinese villages thousands of years ago.

To provide a reasonable chance of success at experimentally inducing substantial variation in both culturally transmitted beliefs and culturally transmitted preferences, we focus on two well studied cultures thought to be quite different in both respects, the US and China. These two cultures have served as central examples in the psychological literature on Eastern vs. Western cultures (Nisbett et al., 2001; Nisbett, 2010) as well as in the expansive transdisciplinary literature on collectivism vs. individualism (Greif, 1994; Triandis, 1995, 2001; Oyserman, Coon and Kemmelmeier, 2002; Oyserman and Lee, 2008; Greif and Tabellini, 2010; Gorodnichenko and Roland, 2011, 2017; Heine, 2015; Greif and Tabellini, 2017; Hajikhameneh and Kimbrough, 2019).[4]

Measuring the impact of culture on cooperation is challenging because it requires, at a minimum, the ability to measure both cooperation and culture together with credible exogenous variation in culture. While there has been much progress in the measurement of societal cooperation and culture, concerns about plausible identification of the relationship between culture and cooperation still prevail. Simply comparing cooperative behavior across two cultures is obviously inadequate because any two cultures differ along many observable and unobservable dimensions, such as institutions, social customs and patterns of interaction, and controlling for variation in these dimensions is often impossible. A similar issue may arise when studying people of varying cultural backgrounds in the same social and economic environment, as is the case with the (second-generation) immigrant studies that have become a workhorse in cultural economics literature. Again, it is plausible that immigrants or their country-of-residence experiences and constraints differ along various unobserved dimensions. In both cases, sources of endogeneity such as omitted variable bias constitute a severe impediment to any attempt at isolating the impact of culture on cooperation.

To tackle this issue, we depart from much of the economics literature on culture and adopt the view that culture is one aspect of an individual's multi-faceted self-concept, or identity (Akerlof and Kranton, 2000, 2002, 2005, 2010). In particular, we assume that behavior at any moment is primarily influenced by the most accessible, activated or context-appropriate cultural identity (Higgins, 1996; Peng and Knowles, 2003; Oyserman and Lee, 2008; Akerlof and Kranton, 2010).This allows us to create the necessary exogenous variation in the influence of culture by experimentally varying the salience of contextual factors which have been shown to activate specific cultural identities among multi-cultural participants (Shih, Pittinsky and Ambady, 1999; LeBoeuf, Shafir and Bayuk, 2010; Benjamin et al., 2010).

Using participants with ties to both US and Chinese culture, i.e., foreign-born (first-generation) and US-born (second-generation) Chinese immigrants studying at a large public university in the US, we examine cooperation in a one-shot anonymous Prisoners' Dilemma (PD) game. This game offers a precise measure of cooperation and captures the type of arms-length non-kin cooperation that economists have singled out as particularly important for economic activity. We randomize participants into different treatments along two dimensions: i) whether US or Chinese cultural identity is salient; and ii) the capacity for culture to determine beliefs.

To manipulate the first dimension (cultural identity salience), we use a simple survey-based instrument borrowed from closely related previous literature (LeBoeuf, Shafir and Bayuk, 2010; Peng and Knowles, 2003). Specifically, we vary questions about Chinese or US cultural events and past experiences with Chinese or US culture. This technique ("social priming") has been widely used in cultural psychology (for an overview, see Matsumoto, 2001; Nisbett, 2010) as well in sev-

[4]China is among the group of countries that are most culturally distant to the US (Muthukrishna et al., 2020). Moreover, on the 0 to 100 Hofstede scale of individualism, with 100 indicating a fully individualist culture and 0 indicating a fully collectivist culture, China scores 20 while the US scores 91 (Hofstede, Hofstede and Minkov, 2010).

eral highly regarded studies in the identity economics literature (e.g., Benjamin et al., 2010; Chen et al., 2014; Cohn, Fehr and Maréchal, 2014; Cohn, Maréchal and Noll, 2015), but is largely absent from economists' study of culture. Although priming studies lost favor and some credibility during the early 2010s replication crisis in psychology, putting aside some of the more outlandish claims an emerging consensus supported by a growing body of evidence holds that priming effects *do* exist and *do* replicate (Molden, 2014; Weingarten et al., 2016; Payne, Brown-Iannuzzi and Loersch, 2016; Cohn and Maréchal, 2016; Lodder et al., 2019; Chivers, 2019). Still, as a nod to replicability concerns we use priming instruments taken essentially verbatim from closely-related research. We vary the second dimension by providing participants information about previous-session cooperation rates in some treatments, but not in others.[5] The idea behind providing concrete and relevant information about cooperation rates is to anchor participants' beliefs about others' PD behavior, thereby constraining culture's effect on beliefs while leaving culture's effect on preferences totally unconstrained.

The major strength of constructing a pool of bi-cultural participants and then manipulating the salience of one of their two cultural identities is that it allows us to circumvent the inherent identification problems associated with culture discussed above. Randomizing which cultural identity is temporarily salient among a pool of people sharing both of the potentially activated identities should ensure that observable and unobservable characteristics are on average the same across the cultural identities we study. Consequently, any difference in cooperation rates we observe plausibly reflects the marginal effect of culture on cooperation. In addition, our design allows us to assess culture's influence on cooperation rates through each of its two defining channels separately. Our treatments providing directly relevant and unbiased information about previous-session cooperation rates should substantially weaken culture's influence on beliefs, leaving preferences as the primary channel through which culture can affect cooperation. One justification is that culture may bring to mind stories, narratives and particular past interactions which play the role of information in its absence, but which are less influential when directly relevant actual information is available. In contrast, treatments without such information provision leave culture's effects on beliefs and preferences totally unconstrained. In these treatments, variation in cooperation rates across cultural identities should therefore reflect culture's effects on beliefs and on preferences. Consequently, the difference in cooperation rates within each salient cultural identity across our information conditions should be attributable to culture's effect on beliefs, as the weakening or, ideally, removal of the beliefs channel is the primary difference between our information conditions. Separately identifying culture's influences through its two defining channels and providing causal evidence on the relative strength of each is the primary contribution of our study.

To formally illustrate how preferences and beliefs could interact to affect stranger cooperation, after presenting our main results we provide a simple framework modeling our particular one-shot anonymous PD game setting. The framework builds on insights from Identity Economics (Akerlof and Kranton, 2010) and from research on social norms (Bicchieri, 2006) to transform the PD into a Bayesian game. Specifically, we assume the degree to which individuals care about identity-prescribed ideals is private information.[6] In this framework, we show that the results of

[5] We vary these dimensions only across sessions in order to minimize the possibility that participants realize there are different treatments. That is to say, in a particular session either all participants' US cultural identity or all participants' Chinese cultural identity is made salient and, in that same session, either all participants receive concrete information about previous-session cooperation rates or no participants receive such information.

[6] We could assume instead that the culturally prescribed ideal action itself is private information. In our simple two-action setting, this would make little difference.

our experiment are consistent with particular Bayesian Nash equilibria of the transformed PD and document that beliefs and preferences can have separate and substantial effects on equilibrium cooperation rates.

As a preview of our results, in our treatments without information provision where culture can affect both preferences and beliefs, we find that US cultural identity substantially *increases* cooperation with anonymous strangers in the PD. That is to say, increasing the salience of US cultural identity increases the proportion of participants choosing to cooperate by about 23 percentage points relative to increasing the salience of Chinese cultural identity, which yields a cooperation rate of about 42 percent.[7] The effect size accords well with similar economics studies manipulating social cues (e.g., Ellingsen et al., 2012; Engel and Rand, 2014; Goerg, Rand and Walkowitz, 2019). When we substantially weaken the beliefs channel by providing previous-session cooperation-rate information, we find the exact opposite. There, increasing the salience of US cultural identity induces only 41 percent of participants to cooperate, which is substantially lower than the 59 percent cooperation rate among participants whose Chinese cultural identity is salient in the same information condition. The comparison between cooperation rates across our information conditions suggests that culture operates as strongly, if not more strongly, through the beliefs channel than through the preferences channel. This is consistent with the widespread notion that Chinese culture is associated with cooperative preferences but pessimistic beliefs about strangers and that the opposite is true for US culture: less emphasis on cooperative preferences but optimistic beliefs about stranger cooperation (e.g., Yamagishi, Cook and Watabe, 1998; Alesina and Giuliano, 2015; Gächter, Schulz and Thoeni, 2017; Enke, 2019).[8]

Our paper makes several contributions. First of all, we provide novel causal evidence on the relative importance of culture's two defining channels of transmission, beliefs and preferences, in determining stranger cooperation.[9] The importance of this contribution lies in the notion that correcting or, more nefariously, exacerbating miscalibrated beliefs may be more easily accomplished, using more straightforward policy measures, than altering culturally transmitted preferences or values (see e.g., Bursztyn, González and Yanagizawa-Drott, 2020).

Second, our study complements a growing body of empirical work showing that culture matters for various economic outcomes (for an overview, see for example, Alesina and Giuliano, 2015). Most of this literature has applied an "epidemiological approach" that exploits exogenous variation in inherited beliefs and values with respect to second-generation immigrants' experiences in their countries of residence to provide causal evidence of culture on outcomes (e.g., Fernández and Fogli, 2006; Giuliano, 2007; Fernández and Fogli, 2009; Algan and Cahuc, 2010; Luttmer and Singhal, 2011). We depart from this approach by providing complementary causal evidence that

[7] We validate the manipulation of identity salience using tools from natural language processing to analyze participants' free-form responses in our survey instrument and show that the manipulation activated country-specific cultural values, practices, and thoughts, as intended. This suggests that our survey instrument temporarily rendered US or Chinese culture identity more salient.

[8] For example, Enke (2019) shows that societies with loose kinship ties are characterized by more cooperation and trust among strangers than tightly-knit societies, which can be accounted for by individual belief systems. Relatedly, cross-cultural research suggests that greater market integration, generally more characteristic of individualism than collectivism, is associated with more cooperative and fair behavior in anonymous transactions (Henrich et al., 2001, 2005, 2010).

[9] Bigoni et al. (2019) provide intriguing correlational evidence. In a repeated public goods game experiment, they find that students from the north and south of Italy exhibit similar preferences for conditional cooperation. In a second experiment, they find that students from both the north and from the south believe northern Italians will contribute more in a similar game. They argue that their findings suggest differences in (conditional) cooperation between the north and south of Italy are explained by differences in beliefs.

gets as close as possible to the ideal situation of observing the *same* individual at the *same point in time* acting under the influence of two different cultures. This allows us to sidestep most concerns about endogeneity due to, e.g., unobserved heterogeneity in immigrant-group experiences (see Fernández (2008, 2011) and Alesina and Giuliano (2015) for a more thorough discussion of these concerns), and permits an examination of the causal impact of cultural identity on cooperation that is precise enough to plausibly identify its effects through separate channels, beliefs and preferences. Our design allows us to do this without relying on elicited beliefs, but rather, on the behavioral implications of those beliefs. This is also an ideal among economists, especially when individuals' beliefs and behavior are simultaneously of interest.[10]

We make another contribution by formally connecting the largely separate literatures on identity and culture in economics.Surprisingly, given the intuitive connection between identity's category-specific prescriptions and the beliefs and values central to economists' definition of culture, to date these bodies of economic research remain largely separate. Connecting these two literatures may benefit both. In particular, the role of cultural identity in coloring beliefs about others features prominently in the culture literature but is absent from the identity economics literature.[11] Consequently, by introducing a simple framework incorporating identity's potential influence on social perception that permits a straightforward way of relating own-values and beliefs about others' values to stranger cooperation while, at the same time, providing evidence about the relative strength of the beliefs channel, our study may open an important new avenue of research for identity economics and advance the literature (Akerlof and Kranton, 2000, 2002, 2005, 2010).[12]

The remainder of the paper proceeds as follows. In the next section we describe our experimental design. In Section 3 we provide empirical results including some robustness exercises. In the penultimate section, we provide a simple theoretical framework to illustrate how culture may affect cooperation in the PD through beliefs and through preferences separately. In the final section we summarize our findings and provide concluding remarks.

2 Experimental Design and Procedures

Our experiment consisted primarily of two phases, an identity-manipulation phase and a game-playing phase. In some treatments, after the Prisoner's Dilemma (PD) was described but before any choices were made participants received information about PD cooperation rates in previous sessions as described in detail below. Overall, we implemented a 2x2 factorial design by varying the cultural identity that was rendered salient (Chinese Identity vs. US Identity) and whether or not participants received (previous-session) PD cooperation rate information. Our design is

[10]Relating elicited beliefs to observed behavior at the individual level is rendered problematic by well known phenomena such as false consensus, social projection or conformity preferences which confound the direction of causation (see, e.g., Butler, Giuliano and Guiso (2015) and the discussion therein). A between-subjects design in which beliefs are elicited in some treatments and behavior is observed in other treatments would be less problematic conceptually but more problematic in terms of sample size and statistical power.

[11]Akerlof and Kranton's model of identity (Akerlof and Kranton, 2000, 2002, 2005, 2010) takes place in a complete and perfect information setting with no role for beliefs formation, while Bénabou and Tirole (2003, 2006*a*,*b*) model identity as beliefs about oneself and leave outside the model how these beliefs might color beliefs about others.

[12]Economists are just beginning to realize the importance of the subconscious automatic beliefs processes associated with identity. In an important recent paper, Bonomi, Gennaioli and Tabellini (2021) model how identification causes voters to slant their beliefs of self and others toward group stereotypes and show how this identity-induced beliefs distortion may exacerbate conflict. The social psychological identity literature upon which identity economics is heavily based, by contrast, features a long history of asking how categorization affects social perception (see the discussion in Butler, 2018).

between-subjects and, moreover, varies each factor only across sessions. Within each session, all participants receive the same cultural identity manipulation and are assigned the same information condition.

Manipulation of Identity Salience In the identity-manipulation phase, participants answer a short questionnaire. The questionnaire begins with filler questions about demographics and student life. The last three items on the questionnaire constitute our manipulation of identity salience. These items follow closely LeBoeuf, Shafir and Bayuk (2010) and vary across treatments (see the Appendix A.2 for the questionnaire). For comparability with previous research, we take the two US-Identity and Chinese-Identity manipulation questions mentioned almost verbatim from their study and add a free-form response question about cultural experiences closely modeled on the instrument in Peng and Knowles (2003). This free-form response question requires participants to actively reflect on cultural practices by asking them to list specific experiences which made them feel comfortable with, or connected to, Chinese or US culture.

The advantage of increasing the salience of US or Chinese cultural identity is that we can unambiguously infer their marginal effect on cooperation as one of the two is always more salient than the other. This is not the case in a neutral condition as we have to assume that both identities exert some influence on behavior in the background and we lack control over its relative strength (see Oyserman and Lee, 2008, for a discussion).[13]

Information Treatments with information about cooperation rates in previous sessions differ only in one aspect from treatments without this information. That is, after the identity-manipulation phase ended, but before decisions were submitted in the PD, participants were presented with a screen containing the following text (where decisions "option L" and "option T" refer to the cooperative outcome): "We have conducted many experimental sessions here in the X-lab over the last year that have involved exactly *this same game*. Participants in these previous experimental sessions were recruited from *the same pool of people* as participants for today's experiment. To give you an idea of how previous participants in the role of the column [row] player played this game, a random selection of 10 participants reveals that X out of these 10 participants chose 'option L' ['option T']." Note that participants were well aware of the composition of the participant pool as they could see the other participants at the beginning of the experiment.

To fill in the "X" for each participant, we drew a sequence of random samples of size ten from previous sessions' PD data. Within each cultural identity treatment, each participant received information derived from a different element in this sequence. To plausibly induce similar or, ideally, identical beliefs across our two identity treatments, we assigned elements of the sequence in parallel to participants across treatments, To illustrate the process, suppose we drew 100 random samples of size ten from previous sessions' PD data which resulted in a sequence $\{X_1, ..., X_{100}\}$

[13]Nonetheless, we also conducted some sessions involving a "Neutral" condition. In this condition, the three "additional" questions and the free-form response questions continued the line of inquiry about student life and we do not provide information about previous-session cooperation rates. The intention of running these sessions was to shed some light on the relative direction of the cultural identity effects, if any, on cooperation. For instance, if we found that the salience of Chinese identity led to more cooperation than the US identity, the Neutral condition could shed light on whether a Chinese cultural identity increases cooperation or whether a US cultural identity reduced cooperation. In fact, we found that the Neutral condition led to a cooperation rate indistinguishable from the cooperation rate in our Chinese Identity treatment where participants were not provided with previous-session cooperation rate information. Among the many plausible interpretations of this pattern is that our bi-cultural students' Chinese identity is typically more salient in their college lives than their US identity.

Figure 1: Prisoners' Dilemma

	Cooperate	Defect
Cooperate	11, 11	0, 12
Defect	12, 0	5, 5

Notes: Entries in the table denote dollar amounts. Rows indicate row players and columns indicate column players. The first entry in a cell refers to a row player's payoff and the second entry refers to a column player's payoff.

of cooperation rates. The first participant in the Chinese Identity treatment with an information-provision phase would receive the information X_1, the second participant would receive information X_2 and so on. The first participant in the US Identity treatment with an information-provision phase would also receive information X_1, while the second US Identity participant would receive information X_2, etc. We drew our sequence of subsamples before conducting any sessions involving an information-provision phase and it was sufficiently long to ensure that we would not exhaust its elements (see Figure A1 for the implemented distribution of information).

This process introduces the same information across the two cultural identities, but different information across individuals, allowing us to estimate the causal impact of culture on *preferences* for cooperation holding (induced) beliefs fixed across cultural identities. This is the primary advantage over the perhaps more straightforward option of transmitting the same information to all participants. It also allows us to truthfully instruct participants that we "randomly selected" the information we were providing without running the risk of selecting unlikely information to transmit (e.g., all previous participants cooperated).

Games and decision measures The game-playing phase was identical across all treatments. It began by randomly assigning participants to pairs to play the one-shot anonymous PD depicted in Figure 1, with payoffs in dollar amounts. Note that the symmetry of the PD permits use of the direct response method, making it particularly appropriate when investigating decisions potentially involving a substantial non-deliberative component such as cooperation. Because our methodology to raise the salience of cultural identity is also likely to rely on a substantial non-deliberative component, direct response becomes even more essential.

We familiarized students with, and quizzed students on, the bi-matrix notation used to describe the PD at the very beginning of each session, i.e., well before the identity-manipulation phase. This allowed us to implement the PD shortly after manipulating the salience of cultural identity, when the effects were most likely to be detectable. After the PD, participants completed a few other tasks (a sequential PD, cheating task, elicitation of risk preferences) and a post-experiment survey to collect some background information. We use these tasks and measures primarily as controls or in the context of robustness exercises, as they are not the focus of our study and the duration of our cultural identity effects is not known with certainty (see Appendix A.1 and A.5 for further details).

Procedures Before beginning the experiment, participants were instructed that there would be several "decision situations" during the experiment and that each separate decision situation would be clearly labeled, but were given no other information about what these decision situ-

ations would entail. Importantly, they were instructed that only one of these situations would be randomly chosen by the computer to determine their entire earnings for the experiment. This design feature is standard and ameliorates concerns about strategic spillovers across tasks.

The experiment was implemented using z-Tree (Fischbacher, 2007) in the Xlab at the University of California, Berkeley. Sessions lasted about 45 minutes and participants earned $16 on average. Because our hypotheses rely on the assumption that participants identify with both a Chinese and US cultural identity, we followed closely related research in cultural psychology (see e.g., Ji, Zhang and Nisbett, 2004) by attempting to recruit only students with a direct connection to a Chinese country (China, Taiwan, Hong Kong, and Singapore), either being born there or having parents born there.

We were hampered in this attempt by policy and information restrictions at the recruiting stage. As a compromise solution, we recruited only participants with a last name appearing on a list of common Chinese last names. This ex-ante screening procedure was, of course, imperfect. As a consequence, we exclude ex-post from our analysis any participant who did not report having at least one parent born in a Chinese country. After these ex-post exclusions, our total sample involves 234 participants.[14] We next present evidence that our sample is balanced on some important observable characteristics.

Balance We collected information on some important demographic characteristics of participants, including age, gender, an indicator for whether the participant reported being an economics major, an indicator for being born in the US, i.e., for being a second-generation immigrant, and background information on parents. Appendix Table A3 presents summary statistics of these variables. They reveal, for example, a roughly equal split between first- and second-generation immigrants in our sample, with 56 percent being foreign-born, and substantial exposure to US culture. On average, foreign-born participants have lived in the US for 6 years and only about 26 percent have lived here less than 2 years. The table also presents the results from non-parametric balance tests, indicating that observables do not differ across treatments (i.e., they neither differ across cultural identities within each information condition nor across information conditions). In addition, Appendix Table A4 provides evidence that observables do not predict treatment status. Together, these tests suggest that randomization into treatments was generally successful and, in particular, alleviate concerns of selection into later treatments.

3 Results

3.1 Manipulation Check

Before proceeding to examine how cultural identity affects cooperation, it is important to examine whether our survey instrument rendered cultural identity more salient. We therefore analyze the extent to which our identity manipulation instrument activated specific cultural cues and thoughts. Specifically, we conduct a text analysis of participants' free-form responses to the question about past cultural experiences and filter for common and rare words by computing each

[14] One might worry that excluding participants ex post who were neither first- nor second-generation Chinese may have introduced an imbalance in the information transmitted, in treatments involving information provision, across cultural identities. However, this was not the case. On average, participants in the Chinese Identity treatment were informed that 50 percent of previous participants cooperated in the PD, while this figure was 48 percent for participants in the US Identity treatment, a difference which was far from significant ($\chi^2(6) = 2.59, p = 0.86$).

Table 1: PD Cooperation Proportions

	Chinese Identity	US Identity
Information provided	0.59	0.41
No information provided	0.42	0.65

Notes: Table reports raw means of cooperation in the four treatment conditions.

word's "term-frequency-inverse-document-frequency" (*tf-idf*). In Appendix Section A.3, we describe and report the results of this textual analysis in more detail. Overall, the analysis revealed that both the US Identity treatment and the Chinese Identity treatment induced responses reflecting country-specific cultural values and practices such as food, festivals, and language, suggesting that our survey instrument achieved the desired goal of making US or Chinese culture temporarily salient.[15]

3.2 Main Result

We now turn to our primary question of how cultural identity affects cooperation as measured by behavior in the PD. In Table 1 we reproduce the raw cooperation proportions, by cultural identity and information condition, mentioned in the introduction. More formally, in Table 2 we report marginal effects estimates from probit models of PD cooperation. In all columns, to account for arbitrary within-session correlations in behavior stemming from, e.g., differences in demographic variables we did not collect or control for, or traits we cannot observe, we cluster standard errors by session.

In column 1, we report a baseline specification controlling only for our experimental factors. There, the coefficient associated with *US Identity* – the excluded category is our Chinese Identity treatment – implies that when both of culture's channels are unconstrained, a US cultural identity induces a significantly higher cooperation rate than a Chinese cultural identity. We can benchmark the magnitude of this treatment effect – 23 percentage points – against other social-cue effects in the existing literature in economics. For example, recent work by Ellingsen et al. (2012) shows that framing a one-shot PD as a community game (vs. a Wall Street game) increases cooperation rates by a similar magnitude, 19 percentage points (see also Engel and Rand, 2014; Goerg, Rand and Walkowitz, 2019, for similar effect sizes).

Substantially weakening the beliefs channel (*Info treatment*) by (ideally) inducing the same set of beliefs across cultural identities increases cooperation rates among participants whose Chinese cultural identity is activated. The estimated marginal effect is sizable, positive and (marginally) statistically significant. The effect of weakening culture's beliefs channel is, however, qualitatively different when US identity is more salient. The marginal effect of the interaction term *US Identity x Info treatment* is negative, highly statistically significant, and large enough in magnitude to imply that weakening culture's beliefs channel substantially *reduces* cooperation rates when a US cultural identity is active. The implied reduction of about 25 percentage points is essentially equal in magnitude to the *increase* in the cooperation rate when a US cultural identity is salient but information about previous-session cooperation rates is not provided, relative to cooperation rates when Chinese identity is salient and neither preferences nor beliefs are constrained. The difference

[15]By contrast, responses in the "Neutral condition" sessions we ran emphasized study and campus-related experiences.

Table 2: Regression: Cooperation rates in the PD

	Dependent variable: PD cooperation				
	(1)	(2)	(3)	(4)	(5)
US Identity	0.235***	0.245***	0.250***	0.295***	0.239***
	(0.090)	(0.087)	(0.086)	(0.104)	(0.092)
Info treatment	0.161*	0.165*	0.172*	0.196**	0.241**
	(0.095)	(0.095)	(0.090)	(0.095)	(0.115)
US Identity x Info treatment	-0.413***	-0.430***	-0.437***	-0.509***	-0.528***
	(0.105)	(0.105)	(0.101)	(0.117)	(0.125)
Individual controls	No	Yes	Yes	Yes	Yes
Cultural controls	No	No	Yes	Yes	Yes
Ancestral roots	No	No	No	No	Yes
Joint sig. controls (p-value)		0.46	0.26	0.69	0.60
Observations	234	234	234	175	130

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Each column reports marginal effects from estimating a Probit model. Robust standard errors clustered by session appear in parentheses. "US Identity" refers to the condition in which US cultural identity is made salient (Chinese Identity is the reference category). "Info treatment" is an indicator referring to the condition in which participants received information about previous-session cooperation rates. We insert a "Neutral condition" indicator variable in all columns. Individual controls include age and indicators for gender, study major (economics), and for being born in the US. Cultural controls include indicators for participants with at least one parent coming from Hong Kong and Taiwan. "Joint sig. controls (p-value)" reports the p-value from a test of joint significance of all controls. Regressions in columns 1-3 use all data. The model in column 4 restricts the sample to participants with both parents born in mainland China and the model in column 5 additionally controls for ancestral roots in the historically wheat-producing regions north of the Yangtze river, as described in the text.

in cooperation rates in the *Info treatment* matches closely the raw data, with a cooperation rate that is about 18 percentage points higher when participants' Chinese identity is salient (compared to US identity).[16] The ability of the beliefs channel to overturn relative PD cooperation rates among the two cultures we study is our primary result.

***Primary result**: Beliefs may play as prominent a role in determining PD cooperation as preferences. When the beliefs channel is unconstrained as in the treatments without information about previous cooperation rates, PD cooperation is about 23 percentage points more likely in the US Identity treatment. When the beliefs channel is substantially weakened, relative cooperation rates are reversed, with PD cooperation being about 18 percentage points more likely in the Chinese Identity treatments.*

Of course, it is possible that our primary result is spurious for various reasons, including that other variables confound it. While our balance tests suggest that demographics do not vary significantly across treatments, one may still worry about their effects on our estimates. We address this concern in column 2 of Table 2 by reporting marginal effects from a probit model in which we explicitly control for participants' observable characteristics including an indicator for being a first- or second-generation immigrant. Our results are unchanged and, reassuringly, demographics (omitted for readability) are jointly non-significant.

We further explore the robustness of our results by addressing the broad concern that the patterns in our data are due to variation in culture across the different Chinese countries in our sample and within China. First, it is conceivable that growing up in Hong Kong under British rule, which may characterize some of our participants' parents' exposure to China, may have induced cultural beliefs or preferences related to stranger cooperation that are closer to those of a Western culture than to those of mainland China; a similar concern may apply to growing up in a country with a more market-based economy like Taiwan. To address these concerns, we control for participants with at least one parent born in Taiwan or Hong Kong (column 3) and restrict attention to participants with both parents born in mainland China (column 4). Both exercises leave our results unaffected.

Second, in column 5 we address possible confounds stemming from cultural variation within mainland China. Even after restricting our sample to participants with both parents born in mainland China, an important cultural confound may remain. Scholars suggest that the historical prevalence of wheat cultivation north of the Yangtze river in mainland China may have shaped a more individualist culture there, while the prevalence of rice cultivation south of the river may have shaped a more collectivist culture (Talhelm et al., 2014; Buggle, 2020; Zhou, Alysandratos and Naef, 2022). We use information on parents' birth cities to construct an indicator for having ancestral roots in the historically wheat-producing regions north of the Yangtze river. Despite the smaller sample due to missing or non-identifiable parental birth place information, our results are robust to restricting our sample to participants having both parents born in mainland China and controlling for ancestral roots.[17]

[16] The difference between Chinese-identity cooperation and US-identity cooperation is given by *Info treatment* - (*US Identity* + *Info treatment* + *US Identity x Info treatment*) = $0.161 - (0.235 + 0.161 - 0.413) = 0.178$.

[17] In the post-experiment survey we asked participants to name the cities in which their mother and father were born. Using these responses, we construct an indicator that takes the value of one if both parents were both north of the Yangtze river and zero if both parents were born south of the Yangtze. If the participant's parents were born on different sides of the Yangtze – this happened for 22 participants – we code our indicator as missing. Alternatively, we randomly assigned these parents to either the north of Yangtze or the south of Yangtze category and re-estimated the model in column 5, which did not change our results (see column 3, Table A6).

In the appendix, we explore and discuss additional dimensions of robustness (see Section A.5). In particular, we examine and discuss alternative specifications to control for cultural variation and for cultural exposure duration (i.e., being a first- or second-generation immigrant). Our results are robust to these robustness exercises.

Summing up, our data reveal that when both of culture's defining channels are unconstrained, a US identity induces more stranger cooperation than a Chinese cultural identity. Substantially weakening the beliefs channel reverses this pattern, suggesting that the two cultures transmit a very different set of beliefs. This is consistent with the commonly held notion that a Chinese identity induces a stronger preference to cooperate in our one-shot anonymous PD setting, but pessimistic beliefs about stranger cooperation, whereas a US identity puts less emphasis on cooperative preferences, but induces more optimistic beliefs about cooperation. Before providing a framework which may help to understand these patterns, we provide some evidence illustrating the differences in culturally transmitted beliefs.

3.3 The Role of Beliefs

Our data strongly suggest that beliefs matter, but are silent, so far, on how. We have not yet exploited the fact that we randomized the specific information participants received about previous-session cooperation in some of our treatments. By restricting attention to these treatments, we can illustrate the role of beliefs along two margins. First, we can use the exogenous variation in induced beliefs to more closely examine the causal impact of different information on behavior, *holding cultural identity fixed*. Second, we can compare cooperation behavior across cultural identities for participants *holding plausibly similar beliefs*. Both exercises allow us, in particular, to illustrate the strong impact of (induced) beliefs on behavior.

We start by examining the broad effects of the information provided on cooperation in the raw data. Pooling observations across cultural identities, as expected, the provided information about others' cooperation is positively related to cooperation ($p = 0.041$). Next, we divide the information provided to participants into three categories: optimistic information (i.e., that more than 5 in 10 previous participants cooperated), neutral information (exactly 5 in 10 previous participants cooperated), and pessimistic information (fewer than 5 in 10 previous participants cooperated). Using these categories, we find that optimistic information generates more cooperation (0.64) than neutral information (0.48) which, in turn, generates more cooperation than pessimistic information (0.39), see Figure A2. The overall decline in cooperation associated with increasingly pessimistic information categories is statistically significant ($p < 0.01$, Jonckheere-Terpstra Test for ordered alternatives). This relationship holds for both the US Identity and the Chinese Identity treatments separately as well, suggesting that different cultural identities do not induce a different reaction to concrete and relevant information. An important takeaway from this analysis is that the *same cultural identity* can result in very different behavior depending on (induced) beliefs, which points to the potentially independent effects of culture's two primary channels.

To get closer to the fundamental question of whether the information we provided substantially weakened culture's direct effects on beliefs, as desired, we turn to another comparison. Within each cultural identity treatment we compare participants across information conditions and show that behavior is similar within cultural identities. For example, we have seen that without receiving information about cooperation rates, an activation of US identity leads to an increase in cooperation. If this is due to optimistic (culturally transmitted) beliefs, we should observe similar cooperation rates among participants whose US identity is salient and who received optimistic

information about cooperation, as defined above. Similarly, focusing on Chinese cultural identity, if the lower cooperation rate among participants who received no information about cooperation rates is partially due to pessimistic culturally transmitted beliefs, then we should see similarly low cooperation rates among participants who received pessimistic information about cooperation rates.

In Appendix Figure A4, we present evidence for these conjectures. Restricting attention to the Chinese Identity treatments and comparing cooperation behavior across the two information conditions, we observe that the cooperation rate among participants who received pessimistic information is essentially identical to participants whose Chinese identity was salient but received no previous-session cooperation information (43 percent vs. 46 percent, Fisher's exact test, $p = 1$). Similarly, turning to the US Identity treatments and comparing the cooperation rate among participants who received optimistic information (55 percent) with the cooperation rate among participants who received no previous-session cooperation information (67 percent) reveals a non-significant difference ($p = 0.42$, Fisher's exact test). Overall, this exercise provides additional evidence of the relative strength of the beliefs channel. It suggests that *but for* differences in culturally transmitted beliefs, increasing the salience of Chinese and US cultural identity would have led to similar cooperation rates.

Both of the exercises directly examining the role of beliefs in detail point to a strong role for beliefs in reversing behavior. In particular, they suggest that pessimistic beliefs about stranger cooperation dampen cooperative behavior, while optimistic beliefs amplify it, as is evident in our treatments without information on previous-session cooperation rates. At the same time, our treatments where participants are provided information about previous-session cooperation rates, which feature a diminished role for the beliefs channel, point to a strong role for culturally transmitted preferences. In the next section we provide a framework which may help to understand how these dual implications could be simultaneously true.

4 Framework

In this section we provide a simple framework that illustrates how culture may operate through both preferences and beliefs to affect PD cooperation. We build on Bicchieri (2006), where uncertainty about whether non-traditional incentives apply may transform a one-shot PD into a Bayesian game (see, Figure A3) in which cooperation is possible in equilibrium. Differently from this previous work, the non-traditional incentive we incorporate is Identity (Akerlof and Kranton, 2010). Our framework extends the identity economics literature in two ways. First, we allow the strength of identity concern, an idiosyncratic identity-concern parameter, to be private information. Second, we assume that (cultural) identity may affect the distribution of these identity-concern parameters as well as, independently, beliefs about this distribution. The purpose of the section is to show that our results are consistent with equilibria of a simple framework incorporating cultural identity, not to argue that this is the only framework consistent with our findings.

We focus on the PD game described in Figure 1, where the numbers indicate money earnings. We transform the game in two ways, by incorporating identity utility and by incorporating private information.

Incorporating Identity To incorporate identity utility, we assume that each player has a currently active cultural identity that prescribes an ideal action, a^I. We follow much of the literature

in assuming that identity utility is a function of the distance between one's chosen action and the ideal action: $dist(a_i, a^I)$. In our two-by-two game setting, we lose little by assuming that the *loss* of identity utility is the product of an idiosyncratic identity-concern parameter (γ_i) and an indicator function taking the value one whenever i does not choose the ideal action

$$I(a_i, a_j; a^I) = -\gamma_i \times \mathbb{1}_{a_i \neq a^I}$$

To account for (traditional) economic preferences, we assume economic utility is represented by own-money earnings.[18] Denote by $x_i(a_i, a_j)$ player i's own-money earnings in the outcome resulting from the pure-strategy pair (a_i, a_j), then

$$u_i(x_i(a_i, a_j), x_j(a_i, a_j)) = x_i(a_i, a_j)$$

Overall utility is just the sum of economic and identity utility

$$U_i(a_i, a_j; a^I) = x_i(a_i, a_j) + I(a_i, a_j; a^I) = x_i(a_i, a_j) - \gamma_i \times \mathbb{1}_{a_i \neq a^I}$$

As a quick example, assume player 1's identity prescribes cooperation, $a^I = C$ and money earnings are as in the PD in our experiment. Player 1's overall utility from cooperating when player 2 cooperates would be $U_1(C, C; C) = 11$ since C involves no loss of identity utility. Defecting, however, does involve a loss of identity utility. Playing D against player 2's C yields overall utility $U_1(D, C; C) = 12 - \gamma_1$. Whether cooperation is a best response to a co-player's action therefore obviously depends on how the identity loss from defecting, γ_1, compares to the monetary gain from defecting. For player 1, cooperation is a best response to defection if $U_1(C, C; C) = 11 \geq U_1(D, C; C) = 12 - \gamma_1$, i.e., if $\gamma_1 \geq 1$. Cooperation is a strictly dominant strategy for player 1 if $\gamma_1 > 5$, since then $U_1(C, D; C) = 0 > U_1(D, D; C) = 5 - \gamma_1$.

Incorporating Private Information To incorporate private information, we assume that each player's identity-concern parameter, $\gamma_i, i = 1, 2$, is private information. For simplicity, we assume that all cultural identities prescribe the same ideal, unconditional cooperation: $a^I = C$. This assumption is primarily for convenience and has the benefit of simplifying notation, eliminating the need for subscripts on ideal action(s).[19]

We assume that players can be one of three types: type-L, type-M and type-H, with associated identity-concern parameters $\gamma_L < \gamma_M < \gamma_H$, respectively. A type-L player is a low identity-concern type who cares about identity sufficiently little that standard economic concerns are always decisive in our PD.[20] Defect is a strictly dominant strategy for type-L players, even after incorporating identity utility. Given the monetary payoffs in our PD, it must be that $\gamma_L < 1$. For the high-concern, type-H, player, identity concerns are always decisive. Cooperate is a type-H

[18] This is a simplifying assumption to focus on the role played by identity. Other common distributional preferences could be incorporated into standard economic utility.

[19] Relaxing the assumption would be straightforward and could incorporate a range of preferences. For example, assuming D as an ideal would be observationally equivalent to assuming $\gamma_i = 0$ in any PD, as then D would always be a strictly dominant strategy. An ideal for conditional cooperation could be captured by allowing ideals to be functions of co-players' actions (e.g., ${a^I}_i = f(a_j)$). In more general games or situations, setting $\gamma_i = 0$ could also capture the possibility highlighted by the High School "burnouts" in Akerlof and Kranton (2002) of particular identities prescribing no ideal action at all.

[20] This, of course, depends on the monetary payoffs chosen for our PD. For this characterization to generally be true, the only identity-concern parameter feasible would be zero

player's strictly dominant strategy. This imposes the condition $\gamma_H > 5$. Finally, Type-M players do not have a strictly dominant strategy in the identity-transformed PD, implying that their best responses depend on beliefs about their co-player's action. The restriction implied by this characterization is $1 \leq \gamma_M \leq 5$.

For the remainder of this section, we choose specific values for players' identity concern parameters which meet the restrictions above. In particular, we fix $\gamma_L = 0, \gamma_M = 3$ and $\gamma_H = 6$.

Culture's two channels Culture affects behavior through its two defining channels, either by altering the distribution of types (preference channel) or by coloring beliefs about the type distribution. The preference channel embodies the plausible assumption that individuals may care about some specific identities more than others. Activating an identity that is particularly important to a lot of people might, for instance, shift the identity-concern parameter distribution toward type-H players and away from type-L players, compared to an identity that is only mildly important to many people.[21] The beliefs channel reflects the conjecture that making a specific identity salient may bring to mind particular narratives or subsets of past interactions that other identities would not. This selective recall may, in turn, color beliefs. Indeed, this is one interpretation of how our experimental cultural identity manipulation operates.

In our analysis below, we assume that culture's two channels may operate independently. So, for instance, making a particular culture salient may affect beliefs about the type distribution irrespective of whether it affects the actual type distribution. An implication of this assumption is that subjective type-distribution beliefs need not be accurate nor common. There may be culturally induced heterogeneity in type distribution beliefs.

Bayesian Nash Equilibria To illustrate how cultural variation in beliefs and preferences may affect cooperation rates, we solve for Bayesian Nash equilibria in cases resembling our experimental treatments. To distinguish between preferences and beliefs, we denote by p_K the actual proportion of type-K players in the population and by $\widehat{p_{K_i}}$ player i's subjective belief about this proportion. When it causes no confusion, we will omit the subscript on subjective beliefs. In all of our analyses, we focus particularly on type-M players, since these are the only players whose behavior can vary across equilibria – both type-L and type-H players have strictly dominant strategies.

With information provision The first case to consider mirrors our treatments with information provision. Our interpretation of these treatments is that we substantially weaken culture's beliefs channel and induced a type-distribution belief that is common across both identities studied. Since culture's effect on preferences is unconstrained in these treatments, however, the actual type distributions induced by our two identities may differ. Consequently, variation in behavior across cultural identities can be attributed to variation in preferences (type distributions). For ease of exposition, suppose that we succeeded in inducing a uniform type distribution belief: $\widehat{p_K} = \frac{1}{3}, K = L, M, H$.

Type-M players considering cooperating believe that with probability $\frac{1}{3}$ they will be matched with a type-L player, who must defect and with probability $\frac{1}{3}$ they will be matched with a type-H

[21] This possibility is most clearly illustrated by nested identities. For example, many individuals are willing to spend a lot of money and exert a lot of effort to show loyalty to their alma mater's football team. However, when their own team has been eliminated from the playoffs, so that the currently salient identity becomes the football conference to which one's team belongs, it would not be surprising to find that fewer individuals would be willing to spend time and money to live up to the ideal of showing loyalty.

player, who must cooperate. With the remaining probability, they will be matched with another type-M player who is cooperating with probability q and defecting with probability $1-q$. The expected utility of cooperating, which involves no loss of identity utility, is

$$\widehat{p_L}0 + \widehat{p_H}11 + \widehat{p_M}(q11 + (1-q)0).$$

Defecting, by contrast, does involve a loss of identity utility. The expected utility from defecting is therefore

$$\widehat{p_L}5 + \widehat{p_H}12 + \widehat{p_M}(q12 + (1-q)5) - \gamma_M.$$

Comparing these expected utilities, assuming that type distribution beliefs are uniform and recalling that we have fixed $\gamma_M = 3$, we see that type-M players may cooperate in equilibrium whenever

$$q \geq \frac{11 - 3\gamma_M}{4} = \frac{11-9}{4} = \frac{1}{2}$$

Type-M players must cooperate in equilibrium when the inequality is strict. By symmetry, a (Bayesian Nash) equilibrium is given by: type-M and type-H players always play C; type-L players always defect. In this equilibrium, the *observed* cooperation rate will be $p_M + p_H$, which may differ by identity.

For illustrative purposes, assume that the type distribution (p_L, p_M, p_H) induced by a US identity is $(0.5, 0.2, 0.3)$, while a Chinese identity induces $(0.4, 0.2, 0.4)$. These distributions embody participants caring more about their Chinese identity: the Chinese type distribution features a lower proportion of low identity-concern (type-L) players and an increased proportion of high identity-concern (type-H) players. These two type distributions would induce the relative cooperation rate pattern we observe in our data. Cooperation would be more prevalent among participants whose Chinese identity is made salient (60 percent) than among participants whose US identity is made salient (50 percent).

Without information provision We next consider our treatments where participants are not provided with information about previous-session cooperation rates. Our interpretation of these "no information" treatments is that both of culture's defining channels – preferences and beliefs – are unconstrained. We assume that culture's preferences channel operates as in the previous case so that a US identity induces the type distribution $(p_L, p_M, p_H) = (0.5, 0.2, 0.3)$ and a Chinese identity induces the type distribution $(0.4, 0.2, 0.4)$. Our task is then to show that there are plausible culturally induced beliefs that can reverse the relative cooperation rates above in equilibrium. This reversal of relative cooperation rates is what we observed in our data across our two information conditions.

To match our intuition, plausible culturally induced beliefs would feature excessive pessimism about type-H (cooperative) players when a Chinese identity is salient and perhaps unwarranted optimism about type-H players when a US identity is salient. For ease of exposition, consider the extreme case where Chinese culture transmits maximally pessimistic beliefs, $\widehat{p_L} = 1$. At the same time, suppose a US culture transmits uniform beliefs, $\widehat{p_K} = \frac{1}{3}, K = L, M, H$. Uniform beliefs would be mildly optimistic about cooperative type-H players – $\frac{1}{3} > 0.3$ – and have the added benefit of allowing us to re-use our expected utility calculations above for US-identity participants.

For participants in the US-identity treatment, we can immediately state that there is an equilibrium in which both type-M and type-H players cooperate. In this equilibrium, the observed cooperation rate would be $p_M + p_H = 0.5$.

Among participants whose Chinese identity is made salient, for type-M players the expected utility of cooperation is higher than the expected utility of defecting whenever

$$\widehat{p_L}0 + \widehat{p_H}11 + \widehat{p_M}(q11 + (1-q)0) \geq \widehat{p_L}5 + \widehat{p_H}12 + \widehat{p_M}(q12 + (1-q)5) - \gamma_M$$

Because the Chinese identity transmits maximally pessimistic beliefs ($\widehat{p_L} = 1$), this inequality boils down to $0 \geq 5 - \gamma_M = 5 - 3$, which is obviously never satisfied. Type-M players therefore always defect in every equilibrium of our transformed PD with these culturally-induced beliefs. Consequently, the observed cooperation rate in equilibrium will be the actual proportion of type-H players, 0.4.

Comparing the equilibrium we identified for the US identity to the equilibria we have just outlined when a Chinese identity is active, we find that observed cooperation rates would match the relative cooperation rate patterns in our data for our no-information condition: higher cooperation in the US identity than the Chinese Identity treatment.

Summary of Framework Exercise The overall message of the exercise in this section is twofold. First of all, we see that the potential influence of the beliefs channel is primarily constrained by the actual population proportion of type-M players. Among those players, beliefs determine whether cooperation occurs or not, while their actual population share determines the effect of their cooperation choices on observed cooperation rates.

The second message is that the reversal of cooperation rates that we observe in our experiment is perfectly consistent with our simple Bayesian game framework in which the strength of identity concern is private information and both the strength of identity-concern and beliefs about the strength of identity-concern may be affected by culture. We have provided simple numerical examples including plausible assumptions about patterns in the beliefs and preferences channels associated with Chinese and US cultural identities that together demonstrate equilibria that closely match the cooperation-rate patterns in our data.

5 Concluding Remarks

Based on the view that cultural identity is one aspect of individuals' typically multi-faceted self-concept, this paper has taken a step forward in investigating the relative importance of preferences and beliefs in culture's impact on economic behavior. Taking an "identity" perspective, we are able to reveal the causal impact of particular cultural identities on cooperative behavior by borrowing a technique, priming, that temporarily amplifies the relative importance of the "primed" cultural identity over non-primed identities and that has been used repeatedly successfully in the identity economics literature.

Our experimental results illustrate the relative importance of the beliefs channel. We present a simple identity-utility framework, following Akerlof and Kranton (2005) and building on Bicchieri (2006), that allows us to demonstrate the potential interplay between preferences and beliefs in determining PD cooperation. To isolate the relative impact of the beliefs channel, we conducted treatments in which we attempted to constrain the direct effect of culture on beliefs by providing factual information about cooperation rates. Substantially weakening the beliefs channel, our

data suggest that Chinese cultural identity enhanced participants' preference for cooperation, confirming common views about collectivistic cultures. Providing no such information and, thus, giving culture free rein to affect both preferences and beliefs, demonstrates the role of beliefs. In treatments without information provision, activating a US cultural identity induced substantially more cooperation than activating a Chinese cultural identity. Comparing cooperation rates across treatments with and without information supports the view that a US cultural identity typically transmits more optimistic beliefs about others' willingness to cooperate relative to Chinese culture. Together, these findings suggest that culture's effect through the beliefs formation process can be stronger than its effect through preferences as we have formalized in our framework.

Although our findings stem from a particular population of bi-cultural students, they provide some guidance for modeling culture's effects on economic behavior. Specifically, the findings showcase the interplay of culturally-induced beliefs and preferences, implying that beliefs should feature more prominently in models of culture and identity. Indeed, we laid out a framework that integrates beliefs about others' values in a parsimonious way into an identity economics model. While the framework is able to capture common views on the collectivism-individualism cleavage that are mirrored in our data, it may have a more general appeal as it demonstrates when beliefs about others' types (social perception) may matter and when they may not. When identity concern is either very strong or very weak, beliefs about others' cooperation may have little sway over own cooperation; when identity concern is moderate, however, such beliefs may be decisive.

Taking our results at face value, one may wonder how they can be reconciled with key findings in the cultural economics literature documenting long-term persistence of cultural traits and behavior. While evidence on long-term persistence is accumulating, it is important to note that there is a recent debate among prominent scholars on cultural persistence. Fernández (2011, p. 484) argues forcefully that "... a definition of culture that considers [culture] to be slow-moving (see, e.g. Guiso, Sapienza, and Zingales (2006)) is rejected. The speed of cultural change depends on how quickly social beliefs and preferences change over time, which in turn depends on the environment broadly speaking ...." This view is, for example, supported by recent evidence on social norm changes in the wake of the 2016 Presidential Election in the US (Bursztyn, Egorov and Fiorin, 2020), and female labor-force participation in Saudi Arabia (Bursztyn, González and Yanagizawa-Drott, 2020). Similarly, more dynamic environments seem to favor cultural change as traditions and customs are less beneficial for current generations (Giuliano and Nunn, 2021).

Our results contribute support to both sides of this debate. We demonstrate the importance of both an ostensibly slow-moving component – preferences – as well as a cultural component that may move much more quickly, i.e., beliefs. Indeed, providing information about others' cooperation rates indicates that particular cultural beliefs may be readily updated to incorporate new information. Whether and, if so, how quickly, these updated beliefs are subsequently absorbed and transmitted beyond the individual by culture is an important open question.

References

Abramitzky, Ran, Leah Boustan, and Katherine Eriksson. 2020. "Do immigrants assimilate more slowly today than in the past?" *American Economic Review: Insights*, 2(1): 125–141.

Akerlof, George A, and Rachel E Kranton. 2000. "Economics and identity." *Quarterly Journal of Economics*, 115(3): 715–753.

Akerlof, George A, and Rachel E Kranton. 2002. "Identity and schooling: Some lessons for the economics of education." *Journal of Economic Literature*, 40(4): 1167–1201.

Akerlof, George A., and Rachel E. Kranton. 2005. "Identity and the economics of organizations." *Journal of Economic Perspectives*, 19(1): 9–32.

Akerlof, George A., and Rachel E. Kranton. 2010. *Identity Economics: How our identities shape our work, wages, and well-being.* Princeton, New Jersey:Princeton University Press.

Alesina, Alberto, and Paola Giuliano. 2011. "Preferences for redistribution." In *Handbook of social economics*. Vol. 1, 93–131. Elsevier.

Alesina, Alberto, and Paola Giuliano. 2015. "Culture and institutions." *Journal of Economic Literature*, 53(4): 898–944.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn. 2013. "On the origins of gender roles: Women and the plough." *Quarterly Journal of Economics*, 128(2): 469–530.

Algan, Yann, and Pierre Cahuc. 2010. "Inherited trust and growth." *American Economic Review*, 100(5): 2060–2092.

Algan, Yann, and Pierre Cahuc. 2013. "Trust and growth." *Annual Review of Economics*, 5(1): 521–549.

Bénabou, Roland, and Jean Tirole. 2003. "Intrinsic and extrinsic motivation." *Review of Economic Studies*, 70(3): 489–520.

Bénabou, Roland, and Jean Tirole. 2006*a*. "A cognitive theory of identity, dignity, and taboos."

Bénabou, Roland, and Jean Tirole. 2006*b*. "Incentives and prosocial behavior." *American Economic Review*, 96(5): 1652–1678.

Benjamin, Daniel J, James J Choi, A Joshua Strickland, et al. 2010. "Social identity and preferences." *American Economic Review*, 100(4): 1913–1928.

Bicchieri, Cristina. 2006. *The Grammar of Society.* Cambridge University Press.

Bigoni, Maria, Stefania Bortolotti, Marco Casari, and Diego Gambetta. 2019. "At the root of the North–South cooperation gap in Italy: Preferences or beliefs?" *The Economic Journal*, 129(619): 1139–1152.

Binzel, Christine, and Dietmar Fehr. 2013. "Social distance and trust: Experimental evidence from a slum in Cairo." *Journal of Development Economics*, 103: 99–106.

Bisin, Alberto, and Thierry Verdier. 2008. "Cultural transmission." In *The New Palgrave Dictionary of Economics, 2nd edition*. , ed. Steven N Durlauf and Lawrence Blume, 1225–1229. Palgrave Macmillan Basingstoke.

Blau, Francine D, and Lawrence M Kahn. 2007. "Gender and assimilation among Mexican Americans." 57–106, University of Chicago Press.

Blau, Francine D, Lawrence M Kahn, Albert Yung-Hsu Liu, and Kerry L Papps. 2013. "The transmission of women's fertility, human capital, and work orientation across immigrant generations." *Journal of Population Economics*, 26(2): 405–435.

Bonomi, Giampaolo, Nicola Gennaioli, and Guido Tabellini. 2021. "Identity, beliefs, and political conflict." *The Quarterly Journal of Economics*, 136(4): 2371–2411.

Buggle, Johannes C. 2020. "Growing collectivism: Irrigation, group conformity and technological divergence." *Journal of Economic Growth*, 25(2): 147–193.

Bursztyn, Leonardo, Alessandra L González, and David Yanagizawa-Drott. 2020. "Misperceived social norms: Women working outside the home in saudi arabia." *American Economic Review*, 110(10): 2997–3029.

Bursztyn, Leonardo, Georgy Egorov, and Stefano Fiorin. 2020. "From extreme to mainstream: The erosion of social norms." *American economic review*, 110(11): 3522–3548.

Butler, Jeffrey V. 2018. "Identity and the theory of the firm." *Handbook of Behavioral Industrial Organization, Cheltenham and Northampton, MA: Edward Elgar Publishing*, 224–53.

Butler, Jeffrey V, Paola Giuliano, and Luigi Guiso. 2015. "Trust, values, and false consensus." *International Economic Review*, 56(3): 889–915.

Butler, Jeffrey V, Paola Giuliano, and Luigi Guiso. 2016. "The right amount of trust." *Journal of the European Economic Association*, 14(5): 1155–1180.

Butler, Jeffrey V, , Paola Giuliano, and Luigi Guiso. 2016. "Trust and cheating." *Economic Journal*, 126: 1703–1738.

Charness, Gary, Uri Gneezy, and Alex Imas. 2013. "Experimental methods: Eliciting risk preferences." *Journal of Economic Behavior & Organization*, 87: 43–51.

Chen, Yan, Sherry Xin Li, Tracy Xiao Liu, and Margaret Shih. 2014. "Which hat to wear? Impact of natural identities on coordination and cooperation." *Games and Economic Behavior*, 84: 58–86.

Chivers, Tom. 2019. "What's next for psychology's embattled field of social priming." *Nature*, 576(7786): 200–203.

Cohn, Alain, and Michel André Maréchal. 2016. "Priming in economics." *Current Opinion in Psychology*, 12: 17–21.

Cohn, Alain, Ernst Fehr, and Michel André Maréchal. 2014. "Business culture and dishonesty in the banking industry." *Nature*, 516(7529): 86–89.

Cohn, Alain, Michel André Maréchal, and Thomas Noll. 2015. "Bad boys: How criminal identity salience affects rule violation." *Review of Economic Studies*, 82(4): 1289–1308.

Ellingsen, Tore, Magnus Johannesson, Johanna Mollerstrom, and Sara Munkhammar. 2012. "Social framing effects: Preferences or beliefs?" *Games and Economic Behavior*, 76(1): 117–130.

Engel, Christoph, and David G Rand. 2014. "What does "clean" really mean? The implicit framing of decontextualized experiments." *Economics Letters*, 122(3): 386–389.

Enke, Benjamin. 2019. "Kinship, cooperation, and the evolution of moral systems." *Quarterly Journal of Economics*, 134(2): 953–1019.

Fernández, Raquel. 2008. "Culture and economics." In *The New Palgrave Dictionary of Economics, 2nd edition*. , ed. Steven N Durlauf and Lawrence Blume, 333–40. Palgrave Macmillan Basingstoke.

Fernández, Raquel. 2011. "Does culture matter?" In *The Handbook of Social Economics, Vol 1A*. , ed. Jess Benhabib, Matthew O. Jackson and Alberto Bisin, 481–510. North-Holland.

Fernández, Raquel, and Alessandra Fogli. 2006. "Fertility: The role of culture and family experience." *Journal of the European Economic Association*, 4(2-3): 552–561.

Fernández, Raquel, and Alessandra Fogli. 2009. "Culture: An empirical investigation of beliefs, work, and fertility." *American Economic Journal: Macroeconomics*, 1(1): 146–77.

Fischbacher, Urs. 2007. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics*, 10(2): 171–178.

Fischbacher, Urs, and Franziska Foellmi-Heusi. 2013. "Lies in disguise – an experimental study on cheating." *Journal of the European Economic Association*, 11(3): 525–547.

Gächter, Simon, Jonathan Schulz, and Christian Thoeni. 2017. "Cultural origins of cooperation." *mimeo*.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy. 2019. "Text as data." *Journal of Economic Literature*, 57(3): 535–74.

Giuliano, Paola. 2007. "Living arrangements in western europe: Does cultural origin matter?" *Journal of the European Economic Association*, 5(5): 927–952.

Giuliano, Paola, and Nathan Nunn. 2021. "Understanding cultural persistence and change." *The Review of Economic Studies*, 88(4): 1541–1581.

Goerg, Sebastian J, David Rand, and Gari Walkowitz. 2019. "Framing effects in the prisoner's dilemma but not in the dictator game." *Journal of the Economic Science Association*, 1–12.

Gorodnichenko, Yuriy, and Gérard Roland. 2011. "Which dimensions of culture matter for long-run growth?" *American Economic Review*, 101(3): 492–98.

Gorodnichenko, Yuriy, and Gérard Roland. 2017. "Culture, institutions, and the wealth of nations." *Review of Economics and Statistics*, 99(3): 402–416.

Greif, Avner. 1994. "Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies." *Journal of Political Economy*, 102(5): 912–950.

Greif, Avner, and Guido Tabellini. 2010. "Cultural and institutional bifurcation: China and Europe compared." *American Economic Review*, 100(2): 135–40.

Greif, Avner, and Guido Tabellini. 2017. "The clan and the corporation: Sustaining cooperation in China and Europe." *Journal of Comparative Economics*, 45(1): 1–35.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2006. "Does culture affect economic outcomes?" *Journal of Economic Perspectives*, 20(2): 23–48.

Hajikhameneh, Aidin, and Erik O Kimbrough. 2019. "Individualism, collectivism, and trade." *Experimental Economics*, 22: 294–324.

Haslam, S Alexander, and Naomi Ellemers. 2005. "Social identity in industrial and organizational psychology: Concepts, controversies and contributions." *International Review of Industrial and Organizational Psychology*, 20(1): 39–118.

Heine, Steven J. 2015. *Cultural Psychology.* WW Norton & Company.

Henrich, J., J. Ensminger, R. McElreath, A. Barr, C. Barrett, A. Bolyanatz, J. C. Cardenas, M. Gurven, E. Gwako, N. Henrich, C. Lesorogol, F. Marlowe, D. P. Tracer, and J. Ziker. 2010. "Market, religion, community size and the evolution of fairness and punishment." *Science*, 327(5972): 1480–1484.

Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, and R. McElreath. 2001. "In search of Homo economicus: Behavioral experiments in 15 small-scale societies." *American Economic Review*, 91(2): 73–78.

Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, R. McElreath, M. Alvard, A. Barr, J. Ensminger, N. S. Henrich, K. Hill, F. Gil-White, M. Gurven, F. W. Marlowe, J. Q. Patton, and D. Tracer. 2005. "Economic man in cross-cultural perspective: Behavioral experiments in 15 small-scale societies." *Behavioral and Brain Sciences*, 28(6): 795–855.

Higgins, E. Tory. 1996. "Knowledge activation: Accessibility, applicability, and salience." *Social Psychology: Handbook of Basic Principles*, 133–168.

Hofstede, Geert, Gert Jan Hofstede, and Michael Minkov. 2010. *Cultures and organizations: Software of the mind: 3rd edition.* McGraw-Hill Professional.

Hwang, Kwang-kuo. 1987. "Face and favor: The Chinese power game." *American Journal of Sociology*, 92(4): 944–974.

Ji, Li-Jun, Zhiyong Zhang, and Richard E Nisbett. 2004. "Is it culture or is it language? Examination of language effects in cross-cultural research on categorization." *Journal of Personality and Social Psychology*, 87(1): 57.

Knack, Stephen, and Philip Keefer. 1997. "Does social capital have an economic payoff? A cross-country investigation." *Quarterly Journal of Economics*, 112(4): 1251–1288.

Knudsen, Anne Sofie Beck. 2019. "Those who stayed: Individualism, self-selection and cultural change during the age of mass migration." *mimeo*.

LeBoeuf, Robyn A, Eldar Shafir, and Julia Belyavsky Bayuk. 2010. "The conflicting choices of alternating selves." *Organizational behavior and human decision processes*, 111(1): 48–61.

Lodder, Paul, How Hwee Ong, Raoul PPP Grasman, and Jelte M Wicherts. 2019. "A comprehensive meta-analysis of money priming." *Journal of Experimental Psychology: General*, 148(4): 688.

Lowes, Sara, Nathan Nunn, James A Robinson, and Jonathan L Weigel. 2017. "The evolution of culture and institutions: Evidence from the Kuba Kingdom." *Econometrica*, 85(4): 1065–1091.

Luttmer, Erzo FP, and Monica Singhal. 2011. "Culture, context, and the taste for redistribution." *American Economic Journal: Economic Policy*, 3(1): 157–79.

Matsumoto, David. 2001. *The Handbook of Culture and Psychology.* Oxford University Press.

Molden, Daniel C. 2014. "Understanding priming effects in social psychology: An overview and integration." *Social Cognition*, 32(Supplement): 243–249.

Muthukrishna, Michael, Adrian V. Bell, Joseph Henrich, Cameron M. Curtin, Alexander Gedranovich, Jason McInerney, and Braden Thue. 2020. "Beyond Western, Educated, Industrial, Rich, and Democratic (WEIRD) Psychology: Measuring and Mapping Scales of Cultural and Psychological Distance." *Psychological Science*, 31(6): 678–701.

Nisbett, Richard. 2010. *The Geography of Thought: How Asians and Westerners Think Differently... and Why.* Simon and Schuster.

Nisbett, Richard E, Kaiping Peng, Incheol Choi, and Ara Norenzayan. 2001. "Culture and systems of thought: holistic versus analytic cognition." *Psychological Review*, 108(2): 291.

Oyserman, Daphna, and Spike WS Lee. 2008. "Does culture influence what and how we think? Effects of priming individualism and collectivism." *Psychological Bulletin*, 134(2): 311.

Oyserman, Daphna, Heather M Coon, and Markus Kemmelmeier. 2002. "Rethinking individualism and collectivism: evaluation of theoretical assumptions and meta-analyses." *Psychological Bulletin*, 128(1): 3.

Payne, B Keith, Jazmin L Brown-Iannuzzi, and Chris Loersch. 2016. "Replicable effects of primes on human behavior." *Journal of Experimental Psychology: General*, 145(10): 1269.

Peng, Kaiping, and Eric D Knowles. 2003. "Culture, education, and the attribution of physical causality." *Personality and Social Psychology Bulletin*, 29(10): 1272–1284.

Shih, Margaret, Todd L Pittinsky, and Nalini Ambady. 1999. "Stereotype susceptibility: Identity salience and shifts in quantitative performance." *Psychological Science*, 10(1): 80–83.

Tabellini, Guido. 2008. "The scope of cooperation: Values and incentives." *Quarterly Journal of Economics*, 123(3): 905–950.

Tabellini, Guido. 2010. "Culture and institutions: Economic development in the regions of Europe." *Journal of the European Economic Association*, 8(4): 677–716.

Talhelm, Thomas, Xuemin Zhang, Shige Oishi, Chen Shimin, Dechao Duan, Xiaoli Lan, and Shinobu Kitayama. 2014. "Large-scale psychological differences within China explained by rice versus wheat agriculture." *Science*, 344(6184): 603–608.

Triandis, Harry. 1995. *Individualism and Collectivism.* Westview Press.

Triandis, Harry C. 2001. "Individualism-collectivism and personality." *Journal of Personality*, 69(6): 907–924.

Weingarten, Evan, Qijia Chen, Maxwell McAdams, Jessica Yi, Justin Hepler, and Dolores Albarracín. 2016. "From primed concepts to action: A meta-analysis of the behavioral effects of incidentally presented words." *Psychological Bulletin*, 142(5): 472.

Yamagishi, Toshio, Karen S Cook, and Motoki Watabe. 1998. "Uncertainty, trust, and commitment formation in the United States and Japan." *American Journal of Sociology*, 104(1): 165–194.

Zhou, Xiaoyu, Theodoros Alysandratos, and Michael Naef. 2022. "Rice farming and the origins of cooperative behavior." *Economic Journal, forthcoming*.

Appendix – For Online Publication Only

The Causal Effect of Cultural Identity on Cooperation

Jeffrey V. Butler and Dietmar Fehr

A.1 Description of Additional Tasks

Following the Prisoner's Dilemma (PD), we randomly and anonymously re-paired participants to play a sequential version of the PD (SPD). We presented the game in its normal form (see Table A1 below), which allows us to elicit participants' behavior in both player roles. (Using the normal-form version of the SPD was also suggested by the training our participants received on bi-matrix notation.) Participants played the game always first as a row player (second mover) and then as a column player (first mover).

Next, participants were given three dishonesty tasks adapted from Fischbacher and Foellmi-Heusi (2013). For each of these opportunities, participants rolled a 10-sided die in private and reported the result, which determined their own earnings. Across each of these three opportunities we varied who else the die roll affected, pitting own earnings against: i) nobody; ii) another randomly chosen participant in the same session; iii) a well-known charity. The mapping between earnings consequences and the number on which the die landed was randomized across individuals in order to minimize the influence of unintentional cheating (e.g., confusing a "6" for a "9").

After the dishonesty opportunities, we elicited participants' risk attitudes using a multiple-price list procedure (adopted from Benjamin et al. (2010)). The procedure included 18 decisions between a certain payment of \$15 and a binary lottery paying \$30 with probability x and \$0 with probability $(1 - x)$, where x increased from 0.25 in increments of 0.03 (see Charness, Gneezy and Imas (2013) for a recent overview of risk elicitation methods).

Finally, participants answered a short exit questionnaire including socio-economic demographics, hypothetical measures of risk and time preferences and measures for cognitive ability. Additionally, we asked participants about their understanding of the experimental tasks, about the reasoning behind their decisions in the experiment as well as about the purpose of the experiment.

Table A1: Sequential Prisoners' Dilemma

	C	D
CC	\$11, \$11	\$0, \$12
DD	\$12, \$0	\$5, \$5
CD	\$11, \$11	\$5, \$5
DC	\$12, \$0	\$0, \$12

A.2 Survey Instrument

1. What is your age?
2. What year in school are you?
3. *[Chinese Identity only]* What is your ethnicity?
4. *[Chinese Identity only]* From which country did your family originate?
5. What is your major field of study?
6. Do you live on or off campus?
7. How many hours do you spend in a typical week (7 days) on the following activities?
   (a) Attending movies, concerts, sports, or other entertainment events
   (b) Participating in physical exercise, recreational sports, or physically active hobbies
   (c) Socializing with friends
   (d) Spending time with family
8. *[US Identity]* Where do you live? *[Chinese Identity]* Where were your parents born? *[Neutral condition]* During this academic year, what was the average number of hours per night you slept on weeknights?
9. What is your favorite [*US Identity:* American holiday?] [*Chinese Identity:* Chinese holiday?] [*Neutral condition:* place on campus?]
10. Everyday, we have various cultural experiences. In the boxes below, please briefly describe up to 5 incidents, activities or encounters on campus that have happened to you that make you feel [*US Identity:* comfortable with American culture.] [*Neutral condition:* comfortable as a student at UC Berkeley.] [*Chinese Identity* connected to your Chinese culture.] It could include people, events, or objects in society or on campus.

A.3 Analyzing the Survey Instrument

To see whether our survey was successful in increasing the salience of cultural identity, we turn to an analysis of our survey instrument. The last question of our survey instrument to manipulate cultural-identity salience was a free-form response question about cultural experiences. This free-form response question requires participants to actively reflect on cultural practices by asking them to list up to five specific experiences which made them feel comfortable with, or connected to, Chinese or US culture. Participants listed more experiences in our US Identity treatment than in the Chinese Identity treatment (Mann-Whitney test, $p = 0.03$), though the average length of their total response – the number of characters used across all experiences – was about the same.

We can use the responses to this question to analyze the extent to which our survey instrument activated specific cultural cues and thoughts. For this purpose we use two common techniques from natural language processing to quantify the content of responses (see e.g., Gentzkow, Kelly and Taddy (2019), for a review of text processing methods and applications in economics). Specifically, we use "stemming" to reduce complexity of the text and then filter for common and rare words using the "term-frequency-inverse-document-frequency" (*tf-idf*). Stemming replaces words with their root, e.g., "america," "american," and "americans" are replaced by its stem america. The *tf-idf* is a simple statistic that indicates the importance of a word in a collection of words (document). Formally, we can write the *tf-idf* as follows:

$$tf - idf_{ij} = \frac{freq_{ij}}{n} \times \log \frac{N}{\sum_i \mathbb{1}_{freq_{ij}>0}} \tag{1}$$

where $tf - idf_{ij}$ is the *tf-idf* for word i in document j, $freq_{ij}$ is the count of word i in document j, n is the total number of words in document j, N is the total number of documents, and $\sum_i \mathbb{1}_{freq_{ij}>0}$ is the number of documents containing word i. Thus the *tf-idf* is the product of two algorithms: the "term frequency," indicating how often a specific word appears in a document divided by the total number of words in that document and the second term, the "inverse document frequency," which is an adjustment for uniqueness of a used word (i.e., words that appear in several documents are discounted). Accordingly, the *tf-idf* is low either for rare words because their "term frequency" is low or for very common words that appear in many documents as then the "inverse document frequency" will be low. In both cases this means that such words do not reflect distinctive features of a document. Table A2 displays the 10 most frequent words in each treatment. Overall, the analysis revealed that both the US Identity and Chinese Identity manipulation induced responses reflecting country-specific cultural values and practices such as food, festivals, and language, suggesting that our survey instrument achieved the desired goal of making US or Chinese culture temporarily salient. By contrast, responses in the "Neutral condition" sessions we ran emphasized study and campus-related experiences.

Table A2: Ten most frequent words

No Info about coop. rates			Info about coop. rates	
US Identity	Chinese Identity	Neutral	US Identity	Chinese Identity
diversity	chinese	speak	game	china
party	organization	activity	test	host
halloween	festival	afraid	freedom	east asian library
america	speak	bear	open	hear
vote	cantonese	floor	say	mandarin
freedom	family	greet	activity	moon
movie	mandarin	professor	concert	parent
speech	hear	libary	football	autumn
usa	we	lunch	acapella	(asian) bakery
culture	dim sum	paper	door	cantonese

A.4 Summary Statistics and Balance Checks

In Table A3, we present summary statistics of a set of observable characteristics of participants separated by information condition. The table indicates that these observable characteristics are very similar across information conditions. For example, in both conditions the average age is about 20 years and about 19 percent of students are majoring in Economics. Most noteably, we see that the share of participants born in the US and with ancestral roots in mainland China is very similar across conditions as well. To provide statistical support for these observations, we report the results from non-parametric Fisher's exact and Mann-Whitney tests in Table A3. Columns labeled "p-val" present the p-values for differences across cultural identities in each information condition and the last column (labeled "p-val*") displays p-values from comparing covariates across the two information conditions. All comparisons have p-values well-above conventional levels of significance ($p > 0.1$), with one exception. The share of students with Taiwanese parents differs across cultural identities in treatments with no information about previous cooperation rates. There is also a slight imbalance of participants with ancestral roots north of Yangtze across information conditions. However, we have to keep in mind that we could not identify the roots for all participants with parents born in mainland China.

To check whether these observables predict treatment status, we estimate the probability of being randomized into one of the treatments on individual characteristics (age, gender, major, etc). Each regression takes the form $y = \alpha + \beta Covariate_i + \epsilon_i$, where $Covariate_i$ is an observed individual characteristic and y is a treatment indicator. Table A4 presents the p-values for each covariate and shows that the randomization is balanced along observable characteristics. All covariates are associated with a p-values well-above conventional levels of significance ($p > 0.1$), except that having at least one Taiwanese parent predicts treatment status in the Chinese Identity setting with no information about cooperation rates. In our main analysis, we always present basic specifications without covariates and compare them to specifications with covariates.

Table A3: Summary statistics

	No Info about coop. rates				Info about coop. rates				
Variable	N	mean	sd	p-val	N	mean	sd	p-val	p-val*
Male	102	0.29	0.46	0.35	132	0.36	0.48	0.47	0.33
Age	102	19.86	2.16	0.81	132	20.17	2.15	0.51	0.12
Major: Economnics	102	0.18	0.38	1.00	132	0.19	0.39	0.86	0.87
Born in US (2nd generation)	102	0.46	0.50	0.12	132	0.42	0.50	0.55	0.60
Years lived in US (if foreign-born)	52	7.04	5.72	0.57	76	5.76	5.91	0.19	0.19
Parents born in mainland China	102	0.79	0.41	0.27	132	0.71	0.45	0.85	0.17
Taiwan	102	0.10	0.30	0.05	132	0.19	0.39	0.83	0.07
Hong Kong	102	0.06	0.24	0.68	132	0.05	0.22	1.00	1.00
Ancestoral roots north of Yangtze	67	0.21	0.41	0.62	76	0.36	0.48	0.23	0.07

Notes: "p-val" refers to p-values for differences across cultural identities in each information condition derived from non-parametric Fisher's exact and Mann-Whitney tests. "p-val*" refers to tests comparing covariates across information conditions (without and with information about previous cooperation rates). "Ancestral roots north of Yangtze" is the share of participants with ancestral roots north of Yangtze among those for whom we can definitely identify their roots. Parents of $n = 22$ participants are born on either side of Yangtze and for $n = 10$ we were not able to identify the roots because either participants did not indicate the city of birth for their parents or it was not possible to locate one of the cities of birth.

Table A4: Treatment Status and Covariates

	No Info about coop. rates			Info about coop. rates	
	US Identity	Chinese Identity	Neutral	US Identity	Chinese Identity
	(1)	(2)	(3)	(4)	(5)
Male	0.36	0.69	0.27	0.91	0.36
Age	0.52	0.84	0.83	0.69	0.45
Major: Economics	0.70	0.46	0.44	0.29	0.32
Born in US (2nd generation)	0.30	0.21	0.22	0.66	0.96
Years lived in US (if foreign-born)	0.39	0.25	0.12	0.15	0.78
Parents born in mainland China	0.96	0.12	0.92	0.62	0.40
Taiwan	0.96	0.05	0.94	0.54	0.25
Hong Kong	0.80	0.46		0.88	0.78
Ancestoral roots north of Yangtze	0.21	0.14	0.72	0.03	0.79
$Prob > F$	0.83	0.25	0.59	0.46	0.68

Notes: Columns 1–5 report p-values from a series of regressions of the form $y = \alpha + \beta Covariate_i + \epsilon_i$, where $Covariate_i$ is the variable listed in the row and y is a treatment indicator. $Prob > F$ is the p-value from an F-test for joint significance of all covariates.

A.5 Additional Robustness Checks

In this section, we present several additional robustness checks validating our results in Section 3.2. We do this along three margins. First, we control for risk preferences, dishonesty, and reciprocity, factors that have been shown to influence cooperation in the PD. Second, we provide alternative specifications for the results on cultural variation and exposure to culture. Third, we examine the possibility that participants differ in what they believe about the experiment and its procedural details.

We start with examining the relationship between cooperation and risk, dishonesty, and reciprocity. Table A5 shows the result. In column 1, we reproduce, for comparison, our main result including covariates from Table 2. In column 2, we control for risk preferences using the risk premium, i.e., the expected return offered by the gamble in excess of the safe option. The risk premium is calculated as the ratio of the difference between the expected value from a gamble and the certain payment. For example, if a participant would choose the gamble of \$30 over the certain payment of \$15 if and only if the winning probability is at least 58 percent, then the risk premium would be $(30 \times 0.58 - 15)/15 = 0.16$. In column 3, we control for participants' dishonesty using a index of their payoffs from the three dice-rolling tasks. To construct this index, we standardize the three dice-rolling outcomes to have zero mean and a standard deviation of one and calculate an equally-weigthed average of the three outcomes. In column 4, we include an indicator for conditional cooperation (i.e., chosing cooperation as a second mover in the SPD if the first mover cooperates and defection otherwise). Column 5 include all measures simultaneously. Neither factor greatly affects the estimated effects of our treatments.

Next, we focus on cultural variation and exposure to culture and possible demand effects. We present the results of these exercises in Table A6. Again, the first column reproduces, for comparison, the basic specification with individual covariates from Table 2. In our main analysis, we control for cultural variation across Chinese countries. Alternatively, we can restrict our sample to participants with no ancestral roots in Taiwan and Hong Kong (see column 2). We find qualitatively similar results when we use this alternative specification. In Table 2, we also presented results accounting for the historical prevalence of wheat cultivation north of the Yangtze river in mainland China that may have affected the evolution of cultural traits. While we could not identify the ancestral roots of all participants due to missing and non-identifiable birth locations, we also dropped from our analysis $n = 22$ participants with one parent born on either side of the Yangtze river because in these cases it is not clear how potential cultural differences affected the formation of preferences and beliefs. To include these participants with mixed parents in our analysis, we randomly assign them to either the wheat or rice cultivation region. Column 3 reports the results and, again, shows that our results are robust to this alternative specification.

As another robustness exercise, we consider the possibility that cultural exposure duration (i.e., being a first- or second-generation immigrant) influences our results. While we control for first- and second-generation participants in our main analysis, we present here two additional exercises to address this concern. As a first pass, we use a continuous variable indicating the "years lived in the US" instead of using a binary measure for first- and second-generation participants (column 3). We obtain very similar results as with controlling for whether participants are born in the US or not. In a second exercise, we then test for differences in cultural assimilation by interacting all treatment indicators with an indicator for whether a participant is born in the US. The estimates in column 4 reveal that our results are robust to the inclusion of these interactions. This is consistent with a recent literature on cultural assimilation suggesting that some aspects of cul-

tural assimilation may happen quickly, while other aspects of an immigrant's country-of-origin culture may persist for more than one generations (Blau and Kahn, 2007; Fernández and Fogli, 2009; Luttmer and Singhal, 2011; Alesina and Giuliano, 2011; Blau et al., 2013; Abramitzky, Boustan and Eriksson, 2020), particularly certain customs from collectivist immigrants' home cultures (Knudsen, 2019). It also reflects the notion that some aspects of culture, such as preferences, are slow moving, while other components, such as beliefs, are moving faster.

Finally, we examine possible demand effects that may compromise our estimates. Specifically, we focus on the possibility that our estimates are biased by participants who did not believe our payment procedure (i.e., one randomly determined task) and who identified the purpose of our experiment. In the post-experimental survey, we directly asked participants whether they believed that their responses in the task would matter for their payment. Believing that responses are not tied to payment may result in low-quality decisions and add noise to our outcome measure resulting in biased estimates. Column 6 shows that dropping this group of participants who distrusts our payment procedure does not greatly affect the results, despite the smaller sample. Participants also answered a free-from response question about the purpose of the study. We flag all responses that mention (cultural) background or include any allusion to China and the US (14 participants out of 234) and construct an indicator encompassing all flagged participants, i.e., those who identified the purpose the study. Focusing our analysis on participants who are agnostic about the research hypotheses leaves our results unchanged as well (column 7).

Table A5: Robustness: Cooperation rates in the PD

	Dependent variable: PD cooperation				
	(1)	(2)	(3)	(4)	(5)
US-Identity	0.245***	0.240***	0.245***	0.241***	0.277***
	(0.087)	(0.082)	(0.083)	(0.086)	(0.056)
Info treatment	0.165*	0.157*	0.166*	0.141	0.107
	(0.095)	(0.094)	(0.093)	(0.093)	(0.071)
US-Identity x Info treatment	-0.430***	-0.422***	-0.444***	-0.415***	-0.427***
	(0.105)	(0.101)	(0.102)	(0.104)	(0.078)
Individual controls	Yes	Yes	Yes	Yes	Yes
Experimental controls	No	Risk	Dice	SPD P2	All
Joint sig. controls (p-value)	0.46	0.53	0.40	0.37	0.37
Observations	234	228	234	234	228

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Each column reports marginal effects from estimating a Probit model. Robust standard errors clustered by session appear in parentheses. "US Identity" refers to the condition in which US cultural identity is made salient (Chinese Identity is the reference category). "Info treatment" refers to the condition in which participants received information about previous-session cooperation rates. Individual controls include age and indicators for gender, study major (economics), and for being born in the US. "Joint sig. controls (p-value)" reports the p-value from a test of joint significance of all individual controls. Regression in column 2 controls for a participant's risk premium elicited in a multiple price list. Column 3 controls for dishonesty using a standardized index of the payoffs in the three dice-rolling tasks. Column 4 includes an indicator for conditional cooperation as a second mover in the sequential Prisoner's Dilemma (SPD) and column 5 controls for all measures simultaneously.

Table A6: Robustness: Cooperation rates in the PD

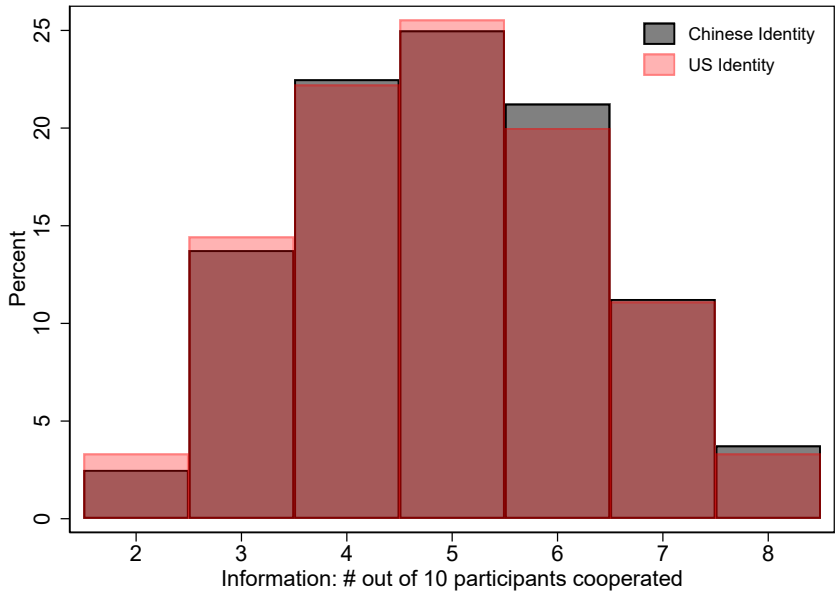
	Dependent variable: PD cooperation						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
US-Identity	0.245***	0.276***	0.223**	0.218**	0.305***	0.292**	0.274***
	(0.087)	(0.106)	(0.094)	(0.092)	(0.114)	(0.114)	(0.089)
Info treatment	0.165*	0.180*	0.176*	0.156	0.142	0.182*	0.180*
	(0.095)	(0.095)	(0.096)	(0.100)	(0.124)	(0.109)	(0.102)
US-Identity x Info treatment	-0.430***	-0.493***	-0.441***	-0.406***	-0.433***	-0.407***	-0.452***
	(0.105)	(0.118)	(0.120)	(0.109)	(0.139)	(0.138)	(0.109)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cultural controls	No	Yes	Yes	No	No	No	No
Joint sig. controls (p-value)	0.46	0.92	0.41	0.53	0.50	0.14	0.54
Observations	234	186	151	231	234	175	220

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Each column reports marginal effects from estimating a Probit model. Robust standard errors clustered by session appear in parentheses. "US Identity" refers to the condition in which US cultural identity is made salient (Chinese Identity is the reference category). "Info treatment" refers to the condition in which participants received information about previous-session cooperation rates. Individual controls include age and indicators for gender, study major (economics), and for being born in the US. Cultural controls include indicators for participants with at least one parent coming from Hong Kong and Taiwan. "Joint sig. controls (p-value)" reports the p-value from a test of joint significance of all controls. Regressions in columns 1, 4, and 5 use all data. The model in column 2 focuses on participants with no ancestral roots in Taiwan and Hong Kong, while the model in column 3 restricts the sample to participants with both parents born in mainland China and simultaneously controls for ancestral roots in the historically wheatproducing regions north of the Yangtze river (randomly assigning mixed parents to either side). Columns 6 and 7 restrict the sample to participants who believed in the payment procedure and who did not identify the purpose of the study, respectively

A.6 Additional Figures and Tables

Figure A1: Implemented Information about Previous Cooperation Rates – US vs. Chinese Identity



Notes: The information "# out of 10 participants cooperated" was presented in neutral terms (see main text). We randomly sampled the information by repeatedly drawing behavior of 10 participants from our previous PD data. We implemented the same information across the two cultural identities. Differences in the distribution are due to excluding ex-post participants with common Chinese last names, but who were neither first- nor second-generation Chinese. The difference is not significant ($\chi^2(6) = 2.59, p = 0.86$).

Figure A2: Provided Information and Cooperation Rates

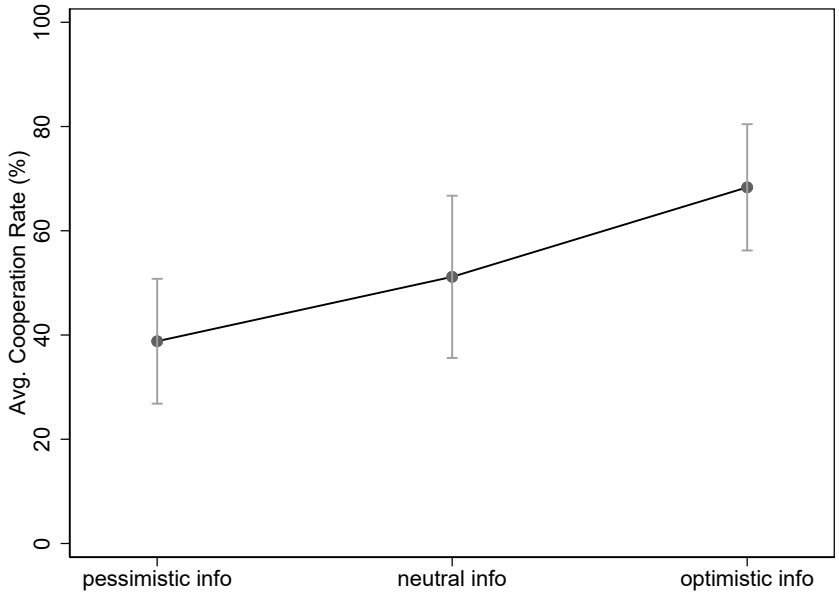
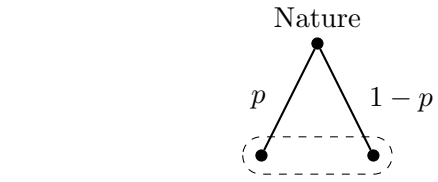


Figure A3: Bayesian Game, adapted from Bicchieri (2006)



	C	D
C	S,S	W,B
D	B,W	T,T

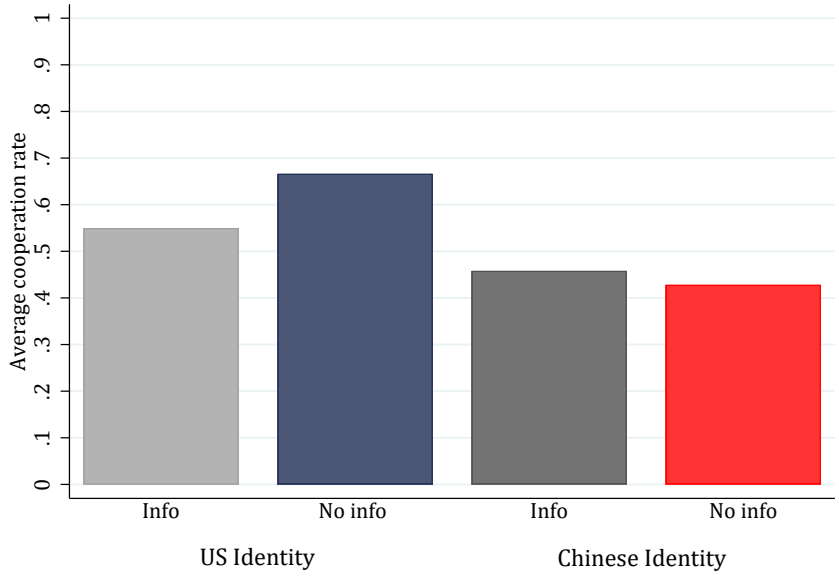
Prisoners' Dilemma

	C	D
C	B,B	W,T
D	T,W	S,S

Coordination Game

Notes: In the figure, B, S, T, and W refer to the Best, Second best, Third best and Worst own-money outcomes, respectively. The bi-matrix on the left defines a Prisoners' Dilemma game, while the bi-matrix on the right defines a coordination game. The essential point that Bicchieri (2006) makes is that incorporating a "social norm" into preferences may transform a Prisoner's Dilemma into a coordination game and that uncertainty over whether a social norm applies may transform the situation into the Bayesian game pictured.

Figure A4: Cooperation Rates across Treatments. Accounting for Implied Beliefs



Notes: "US Identity" and "Chinese Identity" indicate which cultural identity was salient. "No Info" refers to treatments in which participants received no information about previous-session cooperation rates. "Info" here, as discussed in the text, refers to treatments in which participants received information about previous-session cooperation rates that would be consistent with the relative cooperation rate in the "No Info" treatments. For example, the first bar from the left reports the cooperation rate among US Identity treatment participants who exogenously received "optimistic" information about previous-session cooperation rates; the second bar shows the cooperation rate among US Identity treatment who reveived no information about previous-session cooperation rates. Similarly, the two right-most bars present cooperation rates for the exogenously (third bar) and endogenously (fourth bar) induced "pessimistic" information in the Chinese Identity treatments.

A.7 Instructions

A.7.1 General Instructions

Welcome to the study to today. In this study we are interested in understanding individual decision making in strategic and/or risky environment. The study is funded by a large research foundation.

From now on, we ask you to turn off your cell phones. During the experiment you are not allowed to communicate with other participants in any form. If you have questions, please raise your hand. The experimenter will then answer your question in private. Please do not ask questions out loud. If a question is relevant for all participants, we will repeat your question and answer it for all participants. In case you do not follow these rules, we have to exclude you from the experiment.

The study will be completely anonymous. We, the experimenters, will not collect your name, Cal ID or any other identifying information. You have been assigned a participant number and this number will be used throughout the experiment.

The study today consists of several parts. In each part you will be ask to complete different tasks. Before each task you will get instructions about the task and you will be informed how your earnings from this task will be determined. In some tasks these earnings depend on your decision and the decision of another participant in this room, in other tasks your earnings are entirely determined by random chance, and in other task your earnings depend on your decision and random chance.

After completing all tasks, the computer will randomly select a task, which counts towards your final earnings. That is, you will be only paid for one randomly chosen task. However, it is important and in your interest to treat the decisions in each task as if it could be the one that determines your final earnings.

In this study you will get a minimum payment of $5. This $5 is yours. Additionally, you may earn a substantial amount of money from the randomly chosen task. All payments will be made by check at the end of the study today.

A.7.2 Instructions - Normal Form Representation

First part of the study

In the first part of the experiment, you will make multiple decisions affecting your earnings and the earnings of another person in this room. These decisions will be made using a table like the one below, so it is important that you understand what the numbers in the table mean.

Game Table, Example 1

	L	R
T	$1, $2	$3, $4
B	$5, $6	$7, $8

For each decision, you will be assigned a role: either "row player" or "column player." The row player will choose the row of the table and the column player will choose the column. So, for example, if you are assigned to be the "row player" for the game in Example 1 above, you would simply choose between "T" or "B." The column player would choose between "L" and "R." Both

of your choices together would determine one cell of the game table. In this cell are two numbers which describe both of your earnings from the game: ($row player, $column player).

As an example, suppose you are assigned the role "column player." You would be paired with somebody who is assigned the role "row player." You would choose between "L" and "R" without knowing the row player's decision. The row player would choose between "T" and "B" without knowing your decision. Suppose you choose "R" and the row player chooses "B." This means both of your earnings from this game are in the bottom, right-most cell: ($7, $8). Your earnings from the game would be $8, while the row player would earn $7.

Questions: Suppose you are playing the game described in Example 1, above. Suppose you are assigned the role "row player" and you choose "T" and that you are paired with a column player who chooses "L."

1. How much would you, the row player, earn from this game? [Answer: $1]
2. How much would the column player earn from this game? [Answer: $2]

You will also make choices using a slightly more complicated game table. These game tables may have more than two rows and/or more than two columns. An example of such a game table appears below (Example 2).

Even though the table in Example 2 looks more complicated and has more cells, the game is played in exactly the same way as the simpler game of Example 1. You will be assigned one of two roles: "row player" or "column player." The person assigned the row player role chooses the row. The person assigned the column player role chooses the column. The row player's choice together with the column player's choice determines one cell in the table. This cell specifies both the row player's and the column player's earnings: ($row player, $column player).

Game Table, Example 2

	L	R
TT	$1, $2	$3, $4
TB	$1, $3	$4, $2
BT	$3, $2	$5, $2
BB	$5, $6	$7, $8

Questions: Suppose you are playing the game described in Example 2, above. Suppose you are assigned the role "row player" and you choose "BT" and that you are paired with a column player who chooses "L." 1) How much would you, the row player, earn from this game? [Answer: $3] 2) How much would the column player earn from this game? [Answer: $2]

A.7.3 Instructions - Tasks

Instructions Prisoners' Dilemma You will now begin with the first part of this experiment. We will shortly show you the first game table. Depending on your assigned role you have to decide either between the rows or between the columns of the table. If this task is chosen to determine your final earnings, we will randomly match you with another participant in this room. (If you are a row player, we will match you with one of the other column players and if you are column player

we will match you with one of the other row players.) Your decision and this other participant's decision will then determine your earnings from this task.

Instructions Sequential Prisoners' Dilemma You will now begin with the second part of this experiment. In this task you will make a decision in a similar game as before. However, this time we ask you to make a decision for both roles. That is, you will first make a decision in the role of the row player and then in the role of the column player. After all participants made both decisions we will randomly match you to another participant. For each pair of players, we will then randomly determine who will be the row player and who will be the column player. Suppose you are randomly assigned the role of the column player, then the role of the other participant you are matched with is the row player. The earnings in this task are then determined by the decision you made in the role as a column player and by the decision the other participant made as a row player. Because you can end up in either role – row player or column player – it is in your interest to treat the decision in each player role as if it is relevant for your earnings.

Dishonesty Task You will now begin a new part of this experiment. This sections consists of multiple questions (tasks). For each question, your earnings will be determined entirely by chance. If this section is chosen to count towards your earnings, we will randomly select one question from this section and pay you according to your response to that question.

When you have read these instructions, Please click [OK] to proceed to the questions.

Question 1: First, please place the 10-sided die into the cup we have provided you. Next, pick up the cup and place your hand over the top of the cup and shake the cup vigorously for about 5 seconds so that the die tumbles around inside the cup. When five seconds have elapsed, stop shaking the cup and remove your hand from the top of the cup. Please take note of the number on the side of the die facing directly upward and report it below.

The number on the side of the die facing directly upward is:

How this number determines your earnings from this question: [The blanks should be filled in by the computer according to the random permutation that appears in the conversion table] We will convert the number you report into a percentage according to the conversion table below. This percentage determines the percent of $20 that you will earn from this question. For example, if you roll the number [blank], this translates into [blank]% according to the table below, so that your earnings from this question will be [blank]% of $20 = $[blank].

Conversion table: reported number to percentages: [Note: percentages are permuted]

Number	0	1	2	3	4	5	6	7	8	9
Percentage	P0%	P10%	P20%	P30%	P40%	P50%	P60%	P70%	P80%	P90%

Question 2: First, please place the 10-sided die into the cup we have provided you. Next, pick up the cup and place your hand over the top of the cup and shake the cup vigorously for about 5 seconds so that the die tumbles around inside the cup. When five seconds have elapsed, stop shaking the cup and remove your hand from the top of the cup. Please take note of the number on the side of the die facing directly upward and report it below.

The number on the side of the die facing directly upward is:

How this number determines your earnings from this question: [The blanks should be filled in by the computer according to the random permutation that appears in the conversion table] For this question we will pair you with one other randomly selected person in this room. This pairing will be anonymous: you will not learn who the person you have been paired with and, similarly, the other person will not learn who you are. We will convert the number you report into a percentage according to the conversion table below. Suppose this percentage is P%. Your earnings from this question will be P% of \$20, while the other person's earnings will be $(100 - P)$% of \$20. For example, if you roll the number [blank], this translates into [blank]% according to the table below, so that your earnings from this question will be [blank]% of \$20 = \$[blank] and the other person's earnings will be $(100 - P)$% of \$20.

Conversion table: reported number to percentages: [Note: percentages are permuted]

Number	0	1	2	3	4	5	6	7	8	9
Percentage	P0%	P10%	P20%	P30%	P40%	P50%	P60%	P70%	P80%	P90%

Question 3: First, please place the 10-sided die into the cup we have provided you. Next, pick up the cup and place your hand over the top of the cup and shake the cup vigorously for about 5 seconds so that the die tumbles around inside the cup. When five seconds have elapsed, stop shaking the cup and remove your hand from the top of the cup. Please take note of the number on the side of the die facing directly upward and report it below.

The number on the side of the die facing directly upward is:

How this number determines your earnings from this question: [The blanks should be filled in by the computer according to the random permutation that appears in the conversion table] For this question the number you roll will determine you own earnings as well as the size of the donation we make to a well-respected charity, which we will disclose at the end of this experiment. The donation will not be in your name, but we will provide you with a method to verify that we have in fact made the donation. We will convert the number you report into a percentage according to the conversion table below. Suppose this percentage is P%. Your earnings from this question will be P% of \$20, while the amount of money we donate to the charity will be $(100 - P)$% of \$20. For example, if you roll the number [blank], this translates into [blank]% according to the table below, so that your earnings from this question will be [blank]% of \$20 = \$[blank] and the amount we will donate to the charity is (100-P)% of \$20.

Conversion table: reported number to percentages: [Note: percentages are permuted]

Number	0	1	2	3	4	5	6	7	8	9
Percentage	P0%	P10%	P20%	P30%	P40%	P50%	P60%	P70%	P80%	P90%

Risk Elicitation Task You will now begin a new part of this experiment. In this part you will make a series of decisions between one certain option – Option A – and one uncertain option – Option B. Option A will be a fixed and certain payment of \$certain payment. Initially, Option B will be a 100 in 100 chance of \$0 and a 0 in 100 chance of \$winning amount. As you proceed down the rows, Option B will change. The chance of receiving \$winning amount will increase, while the chance of receiving \$0 will decrease. For each row, all you have to do is decide whether you prefer Option A or Option B.

You will have to indicate your preference by checking the corresponding box. Most people begin by preferring Option A and then switch to Option B, so one way to view this task is to determine the best row to switch from Option A to Option B. You should NOT switch back and forth several times between lottery and sure amount.

If this part is chosen to count towards your earnings, we will randomly select one row and pay you according to your decision in that row. For example, suppose your decision the in the selected row was the following: Option A was a certain payment of $4 and Option B was 75 in 100 chance of $0 and a 25 in 100 chance of $8. If your response in this decision was Option A, you would receive $4. If instead your response was Option B then the computer draws a number from 1 to 100 to determine the outcome of the lottery. Note that each of these numbers between 1 and 100 has the same probability of being chosen by the computer. If the randomly drawn number is between 1 and 75 (inclusive) you would receive $0 and if the randomly drawn number is above 75 you would receive $8.

Remember, each decision could be relevant for your payment. So, it is in your interest to treat each decision as if it could be the one that determines your payments.