Cultural differences in responses to real-life and hypothetical trolley problems

Natalie Gold[*]   Andrew M. Colman[†]   Briony D. Pulford[†]

Abstract

Trolley problems have been used in the development of moral theory and the psychological study of moral judgments and behavior. Most of this research has focused on people from the West, with implicit assumptions that moral intuitions should generalize and that moral psychology is universal. However, cultural differences may be associated with differences in moral judgments and behavior. We operationalized a trolley problem in the laboratory, with economic incentives and real-life consequences, and compared British and Chinese samples on moral behavior and judgment. We found that Chinese participants were less willing to sacrifice one person to save five others, and less likely to consider such an action to be right. In a second study using three scenarios, including the standard scenario where lives are threatened by an on-coming train, fewer Chinese than British participants were willing to take action and sacrifice one to save five, and this cultural difference was more pronounced when the consequences were less severe than death.

Keywords: Chinese culture, cultural difference, fatalism, moral decision making, moral judgment, responsibility, Taoism, trolley problem.

1   Introduction

In the classic version of the trolley problem (Thomson, 1985), a runaway trolley threatens to kill five men on the track ahead. A bystander can save the five by switching a lever to divert the trolley on to a side-track where one man will be killed. Moral philosophers have the clear intuition that it would be morally permissible to turn the trolley, and this is a central test case in research on the question of why it is sometimes permissible and sometimes impermissible to harm one person to save many (e.g., Kamm, 2007; Thomson, 1985). The problem has also been used in psychological research into the processes underlying moral judgments (e.g., Greene et al., 2009; Greene, Nystrom, Engell, Darley, & Cohen, 2004; Moore, Clark, & Kane, 2008; Schaich Borg, Hynes, Van Horn, Grafton, & Sinnott-Armstrong, 2006; Waldmann & Dieterich, 2007). Experimenters have found that the average participant's

responses agree with philosophers' intuitions in this case (e.g., Cushman, Young, & Hauser, 2006; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001).

Most trolley research has been done on people from the West, but it is implicitly assumed that intuitions should generalize and that moral psychology is universal. Explicit support for cultural universality has been provided by Hauser, Cushman, Young, Jin, & Mikhail (2007), who reported data from participants in 120 countries and who explored the influence of nationality for every group that was large enough to include. They tested for differences in judgments between a version of the trolley problem in which a passenger on a train can divert it on to a side-track and a footbridge version in which a bystander can stop the train by pushing a man off a footbridge into its path, an action that far fewer respondents endorse. This difference was replicated across Australian, Brazilian, Canadian, Indian, American, and British participants. However, these cultures are all WEIRD (western, educated, industrialized, rich, and democratic, in the terminology of Heinrich, Heine, & Norenzayan, 2010), and if cultural differences exist, they are more likely to be found between cultures that differ more fundamentally. After decades of British colonial rule, even India has a democratic system of government and has evolved largely westernized cultural and educational institutions and practices. In a subsequent study, Abarbanell and Hauser (2010) partially replicated the trolley/footbridge difference in a small-scale, poorly educated agrarian Mayan population.

But there are grounds for suspecting that Chinese culture, in particular, has features that may engender differ-

[*]Philosophy Department, King's College London, Strand, London, WC2R 2LS, UK. Email: Natalie.gold@rocketmail.com.

[†]University of Leicester, UK.

ent responses to trolley problems. Some evidence suggests that East-West cultural differences affect cognitive processes (Nisbett, Peng, Choi, & Norenzayan, 2001), and that these differences lead to differences in judgment and decision-making (Weber & Morris, 2010) and in philosophical intuitions (Machery, Mallon, Nichols, & Stich, 2004; Weinberg, Nichols, & Stich, 2001). A cultural difference that is potentially very relevant to trolley problems is Chinese fatalism—a cluster of beliefs, deeply rooted in Chinese culture, according to which one should allow events to run their natural course without active interference (Bond, 1991; Eberhard, 1966; Kirkland, 2004; Palsane & Lam, 1996). This tends to cause Chinese people, more frequently than others, to attribute life events, including misfortunes, to fate, and to desist from interfering with their progression (Bond, 1991; Palsane, & Lam, 1996).

Chinese and U.S. samples were compared by Moore, Lee, Clark and Conway (2011), using 24 scenarios, translated into Chinese as well as English, "all of which involved sacrificing one person to save multiple others from death" (p. 190). No significant differences were found between the Chinese and U.S. participants in their ratings of how "morally acceptable" different courses of action were to save lives at the cost of others. Nor did Mikhail (2011) find any difference in an experiment run in English, using Chinese immigrants to the US, which asked about the "moral permissibility" of turning the trolley.

Ahlenius and Tännsjö (2012) found that only 52% of Chinese agreed that it is "morally permissible" to flip the switch in the classic trolley problem, as compared to 81% of Americans and 63% of Russians who agree that the agent "should" flip the switch. Their results are difficult to interpret given that the Chinese translation was different from the English and Russian language versions. Ahlenius and Tännsjö claim that the two wordings both express moral judgments and, hence, that they found a cultural difference in moral judgments. The idea that the answer to the question "What ought I to do?" expresses a moral judgment has a long pedigree in philosophy (Hare, 1952). However, in everyday language it is possible to judge that an action is morally permissible and yet think that one should not do it because of other, non-moral considerations. Since moral judgments are not necessarily "all-things-considered" judgments, there can be a gap between what people think is morally right and what they think they should do. Indeed other research on trolley problems has found that moral judgments do not map neatly onto judgments about what the agent should do (Gold, Pulford & Colman, 2014a,b).

Cultural differences might manifest themselves in either moral judgments or actions. It is conceivable that people judge an action to be right and yet, when given the choice, they would not do it. For instance, behavior on matters such as vegetarianism and organ donation does not live up to moral attitudes expressed about those behaviors, even among ethicists (Schwitzgebel & Rust, 2013). As well as the possibility, discussed above, that people could judge an action as morally desirable and yet think that they should do something else, there is also the possibility of weakness of will, where people's actions do not correspond to what they think they should do. Tassy, Oullier, Mancini, and Wicker (2013) found discrepancies between judgments ("*Is it acceptable to...*") and predicted choice ("*Would you do...*") in moral dilemmas, with participants' choice of action being more utilitarian than their judgments.

O'Neill and Petrinovich (1998) reported a preliminary comparison of US and Taiwanese responses to the question "What would you do?" in 25 variations of the trolley problem. They investigated the effects of varying six dimensions of the dilemma: whether it involved taking action or inaction, and differences in the numbers of victims involved, their species, their relationship to the responder, their status and whether the victims were in the situation as a result of social agreement, or through no fault of their own. They reported that the action/inaction, species and kinship dimensions explained the most variance in both the US and Taiwanese samples. However, they did not report any direct comparison of the proportions of US and Taiwanese participants who favored sacrificing one to save many in their trolley problems.

Previous experiments on actions in moral dilemmas, including O'Neill and Petrinovich (1998), have asked participants to predict their own behavior (see also Bartels, 2008; Schaich Borg et al., 2006). Such predictions have been shown to be notoriously unreliable (Osberg & Shrauger, 1986; Vallone, Griffin, Lin, & Ross, 1990; Teper, Inzlicht, & Page-Gould, 2011). For obvious reasons, it is impossible to operationalize the classic trolley problem in a real-life laboratory experiment. However, Gold, Pulford, and Colman (2013) found that substituting economic for mortal harms in trolley problems did not affect the pattern of judgments. In the first experiment reported here, we operationalize a version of the trolley problem in which the harms are small but meaningful economic losses, and we compare the actual choice behavior and judgments of British and Chinese samples.
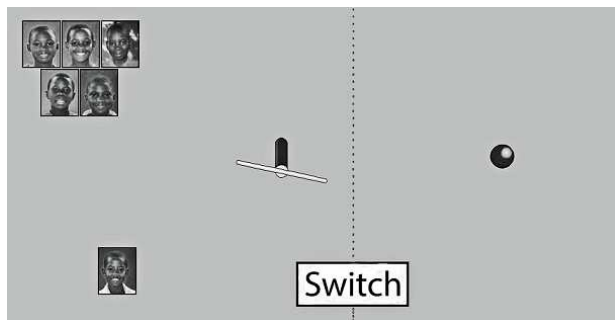
2 Experiment 1

2.1 Method

2.1.1 Participants

We recruited 45 British participants (21 men and 24 women, mean age = 20.73 years, *SD* = 3.73), 61 Chinese participants (7 men and 54 women, mean age = 23.69

Figure 1: Screenshot of ball moving from right to left, on a collision path with a group of five children.

years, SD = 2.20), and 63 other foreign participants who were not native English speakers (17 men and 46 women, mean age = 22.62 years, SD = 4.04). Nationality was classified according to participants' responses to a standard demographic question. We also asked our participants whether they were native English speakers. All the British were native English speakers, none of the Chinese was. We discarded the data of six non-Chinese foreigners who were native English speakers because they were too small a sample to yield meaningful conclusions. Participants were recruited through the University of Leicester's online e-bulletin, which is sent out to students and staff, and were paid £5 ($8) for their participation. They were tested in groups of 15–20.

2.1.2 Procedure and materials

Participants made decisions that influenced the amount of money that we donated to an orphanage in northern Uganda (following a similar protocol of Hsu, Anen, & Quartz, 2008). At the start of the experiment, participants read a brochure from the Canaan Children's Home depicting the children's plight and showing short biographies and photos of some of the orphans, matched for age and gender.

We told the participants that we had endowed each child whose biography they had read with a sum of money that would be enough to supply one meal. That amount was 30p (50c), although we described the payoffs to the participants in terms of meals rather than cash because of the far smaller purchasing power of money in Uganda than in the UK.

Seated at computer monitors, participants then viewed an animation in which a ball moved slowly across the screen toward a group of five children, represented by their photos. On the same screen was a photo of a single child, not in the path of the ball (see Figure 1). On-screen instructions informed the participants that the five children would lose their meals if the ball continued on its current path and hit their photos.

Participants had the option to click on a switch that switched a lever, causing the ball to change direction and head toward the single child, causing that child to lose its meal. What we told the participants was true, and we did, in fact, remove provisionally endowed meals from either five children or one, according to the participants' decisions during the experiment, before sending our donation to the orphanage.

Before participants made their decisions, there were two demonstrations of the animation, one in which the lever was switched and one in which it was not. In the demonstrations, the photos were replaced by blank rectangles, and the number of rectangles in both groups was always five. When participants made their decisions, they had 11 seconds during which they could click the switch before the ball crossed a dotted line in the middle of the screen. The whole animation took 17.5 seconds. For those participants who clicked the switch, their decision times were recorded. Decisions were irreversible and participants knew this in advance. After making their decisions, participants judged, on a 9-point Likert scale from 1 (*Definitely wrong*) through 5 (*Neutral*) to 9 (*Definitely right*), "How wrong or right was it to switch the lever?"

2.2 Results

2.2.1 Decisions

We found that 80.00% of British participants but only 49.18% of Chinese participants clicked the switch. These proportions differ significantly, $\chi^2(1, 106) = 10.47$, $p = .001$, Cohen's effect size $w = 0.31$ (medium). We can control for any effects of characteristics that are specific to those who study abroad, including not being a native English speaker, by using the group of participants who came from abroad but were not Chinese and who did not speak English as their native language. We found that 69.35% of that group clicked the switch, which is significantly different from the Chinese, $\chi^2(1, 123) = 5.19$, $p = .023$. In contrast, there is no significant difference between that group and the British, $\chi^2(1, 108) = 1.84$, $p = .17$.

However, a higher proportion of the foreign participants (both Chinese and non-Chinese) were female, and their mean age was also higher than the British. In order to take account of these confounds, we ran a logistic regression analysis. We found that only the dummy variable for being Chinese was significant, with Chinese participants clicking the switch less frequently than British. The model is shown in Table 1.

When participants clicked the switch, response times were not significantly affected by nationality, gender, or age. Our regression model is shown in Table 2.

Table 1: Logistic regression of clicked (clicked =1, did not click = 0) with nationality (dummy variables for Chinese, and other foreign participants who did not speak English as a native language), gender (male = 0, female = 1), and age.

	B	SE	Wald	p	Exp(B)
Chinese	1.65	0.52	10.07	.002	0.19
Other foreign	−0.76	0.49	2.43	.119	0.47
Gender	0.34	0.41	0.69	.407	1.41
Age	0.03	0.05	0.39	.531	1.03
(Constant)	0.52	1.20	0.19	.664	1.69

Table 2: Multiple regression of decision response time (in ms) on nationality (dummy variables for Chinese, and other foreign participants who did not speak English as a native language), gender (male = 1, female = 0), and age, for participants who clicked. (Unstandardized coefficients are B; standaridze are β.)

	B	SE	β	t	p
Chinese	458.89	632.81	0.09	0.73	.470
Other foreign	860.70	538.58	0.18	1.60	.113
Gender	87.51	523.63	0.02	0.17	.868
Age	54.77	62.62	0.09	0.88	.384
(Constant)	4497.12	1373.37		3.28	.001

2.2.2 Judgments

A multiple regression analysis reveals that Chinese participants gave lower wrong-right judgments, judging the action to be less right, and older participants gave higher ratings. There was no significant effect of gender or of being a non-native English speaker who is not Chinese (see Table 3).

A striking difference is that 31.15% of the Chinese sample judged switching the lever *Neutral*, the modal Chinese judgment, but only 4.44% of the British sample did so. This is consistent with previous findings that Chinese and other East Asians are more likely to use the midpoint of a rating scale than Westerners (e.g., Chuangsheng, Shin-Ying, & Stevenson, 1995). Research has suggested that omitting midpoint response data provides an indication of what would have been obtained if there had been no midpoint on the scale (Schuman & Presser, 1981). If we omit the *Neutral* responses and group the participants according to whether their judgment indicated that switching the lever was wrong (0–4) or right (6–9), then we find that 69.77% of British but only 47.62% of Chinese participants

Table 3: Multiple regression of wrong-right judgments on nationality (dummy variables for Chinese, and other foreign participants who did not speak English as a native language), gender (male = 1, female = 0), and age.

	B	SE	β	t	p
Chinese	−1.44	0.55	−0.27	−2.62	.010
Other foreign	−0.64	0.51	−0.12	−1.25	.212
Gender	0.00	0.46	0.00	0.01	.996
Age	0.21	0.06	0.29	3.70	.000
(Constant)	1.36	1.24		1.10	.273

Table 4: Logistic regression of dichotomous wrong-right judgments (right = 1, wrong = 0) on nationality (dummy variables for Chinese, and other foreign participants who did not speak English as a native language), gender (male = 1, female = 0), and age.

	B	SE	Wald	p	Exp(B)
Chinese	−1.60	0.59	7.42	.006	0.20
Other foreign	−0.46	0.48	0.92	.337	0.63
Gender	−0.13	0.44	0.08	.776	0.88
Age	0.21	0.08	6.33	.012	1.23
(Constant)	−3.23	1.68	3.69	.055	0.04

judged it right to switch the lever. This difference is significant, $\chi^2(1, 85) = 4.30$, $p = .038$, $w = 0.23$ (small to medium effect).

Accounting for other demographic variables, as above, using a logistic regression analysis, we find that age and being Chinese are both significant predictors of whether the action was judged as right or wrong, with Chinese being less likely to judge that the action was right and older participants being more likely. (Note that the age effect is positive while the effect of being Chinese is negative, so age differences cannot account for other observed cultural differences, as the Chinese participants were older.) See Table 4 for the model. The percentages in the dichotomous wrong-right judgments of the British and Chinese are strikingly similar to the percentages of each nationality who actually switched the lever. A McNemar test reveals no significant difference between these two dichotomous distributions, $\chi^2(N = 85) = 0.15$, $p = .70$.

3 Experiment 2

Several criticisms could be made about the results of Experiment 1, which we attempt to control for in Experiment

2. Firstly, the Chinese participants may have understood the scenario less clearly even than the other foreign participants, if English is more difficult for Chinese than for other nationalities. Therefore in Experiment 2 we had British and Chinese participants each complete the experiment in their own native language.

Secondly, the decision in Experiment 1 was taken in an 11-second time frame, and it is possible that the Chinese participants may just be slower to take moral decisions, which would have resulted in the switch not being clicked in time and the decision being recorded as being pro the status quo. To remedy this we used a written scenario, with no time limit, and asked participants what they "would do" if they were in the experiment. This is the wording that has been used by previous researchers so, in order to check our results against those of O'Neill and Petrinovich (1998), we also included the classic trolley problem where lives are at stake from an oncoming train.

Thirdly, we wanted to eliminate the possibility that there was something about the task used in Experiment 1 that interacted with culture and produced different patterns of responses, so we also included a more neutral gameshow scenario where money is at stake. This allows us to control for the possibility that we found cultural differences in Experiment 1 only because of our use of Ugandan orphans, or because the British were more sceptical about the details of the situation and were more likely to believe that all children would get fed fairly by the orphanage with the money provided. Comparing the two scenarios that involve economic harms with the standard trolley problem also allows us to check whether the cultural difference we found in Experiment 1 was related to our substitution of economic losses for mortal harms.

In order to investigate the causes of any cultural differences, we asked participants about their attitude to fate and administered a Locus of Control scale (IPIP; Goldberg, 1999), and we included an open-ended text box where participants could state the reason for their choice of action in order to discover the thinking behind the decisions.

3.1 Method

3.1.1 Participants

We recruited 55 British participants (15 men and 40 women, mean age = 21.11 years, SD = 2.48) and 45 Chinese participants (13 men and 32 women, mean age = 23.73 years, SD = 2.53) in the age range of 18–30 years. All the British were native English speakers, none of the Chinese was. Participants were recruited through the University of Leicester's online e-bulletin, which is sent out to students and staff.

3.1.2 Translation

The English language version of all materials was translated into Simplified Chinese by a native Chinese translator, then back-translated by a different Chinese translator, and the differences reconciled by a third native Chinese speaker who was doing a PhD in philosophy in the UK, to make the Chinese version correspond as closely as possible to the original English.

3.1.3 Procedure and materials

Participants completed the study on-line in their own time via a SurveyGizmo site. After reading the consent information, they read three scenarios in either English or Chinese. The Orphan scenario was a written description of the laboratory experiment that we reported as Experiment 1 and also contained an image of the screen as shown in Figure 1:

> Imagine the following scenario: You volunteer to take part in a psychology experiment. You are given a booklet to read with information about an orphanage called the Canaan Children's Home in Uganda. This contains photos and a short biography about some of the poor children that live there. Then you are told that the choices that you make in the experiment will influence real amounts of money that the experimenters will donate to the orphans. The experimenters tell you that they have endowed each child with a sum of money that would be enough to supply one meal.
>
> Then you are seated at a computer and watch an animation in which a ball moves slowly across the screen toward a group of five children, represented by their photos. On the same screen is a photo of a single child, not in the path of the ball (see Figure below). You are told that the five children will lose their meals if the ball continues on its current path and hits their photos.
>
> You are given the option to click on a switch that moves a lever, causing the ball to change direction and head toward the single child, causing that child to lose its meal.
>
> The direction of the ball can be changed by switching the lever at any time until the ball passes the dotted line. You have 11 seconds during which time you can click the switch to move the lever before the ball crosses the dotted line.

After reading the scenarios participants were asked:
(1) Would you click the switch to move the lever? (Yes/No).

(2) Now please indicate how wrong or how right you think it would be to switch the lever: (1 *Definitely wrong* to 7 *Definitely right*).

(3) Is it morally wrong for you to switch the lever? (Yes/No).

(4) Please explain why you decided to move or not move the lever: (open ended text box).

The Train scenario read:

> Imagine the following scenario: You are taking your daily walk near the train tracks when you notice that the train that is approaching is out of control. You see what has happened: the driver of the train saw five men working on the track ahead and slammed on the brakes, but the brakes failed and the driver fainted. The train is moving so fast that anyone it hits will die immediately. There are five people working on the main track. It is obvious that they will not be able to get off the track in time and, if nothing is done, they will be killed.
>
> The track has a side-track leading off to the left. You are standing next to a lever. If you pull the lever, that will turn the train onto the side track and the five people on the main track will not die. But a person is working on the side track. If the train goes onto the side track, then the person on the side track will die. You are aware of all these facts.
>
> Thus, you can pull the lever, in which case the one person will die but the five people will not; or you can refrain from pulling the lever, in which case the five people will die but the one person will not.

After reading the scenario, participants were asked:
(1) Would you pull the lever? (Yes/No).

Then they received questions (2)–(4) from the Orphan scenario but *switch the lever* was replaced with *pull the lever*.

The Gameshow scenario read:

> Imagine the following scenario: You are a member of the studio audience watching a game show. Five contestants have each earned £100 prize money by answering questions over several rounds, and their tokens are nearing the winning side of the game board. A ball is suddenly released and is rolling towards the tokens of the five contestants and, if nothing is done, they will be knocked out of the game and lose their prize money.
>
> You see that a button on your armrest has just lit up to indicate that you have been randomly selected by computer to take part in the show. You

have the option to press the button and knock the ball onto another path. But another contestant, who has also earned £100 prize money, has a token on the new path and will be knocked out of the game and lose his prize money. You are aware of all these facts.

> Thus you can press the button, in which case the one contestant will lose his prize money but the five contestants will not; or you can refrain from pressing the button, in which case the five contestants will lose their prize money but the one contestant will not.

After reading the scenario, participants were asked:
(1) Would you press the button? (Yes/No).

Then they received questions (2)–(4) from the Orphan scenario but *switch the lever* was replaced with *press the button*.

After the scenarios were completed participants filled in demographic information and completed the IPIP 20-item Locus of Control scale and the following three questions on a 1 *Strongly disagree* to 7 *Strongly agree* scale:

(1) I believe it is usually a good idea to allow things to run their natural course

(2) I believe that it is usually best not to try to interfere with the natural course of events
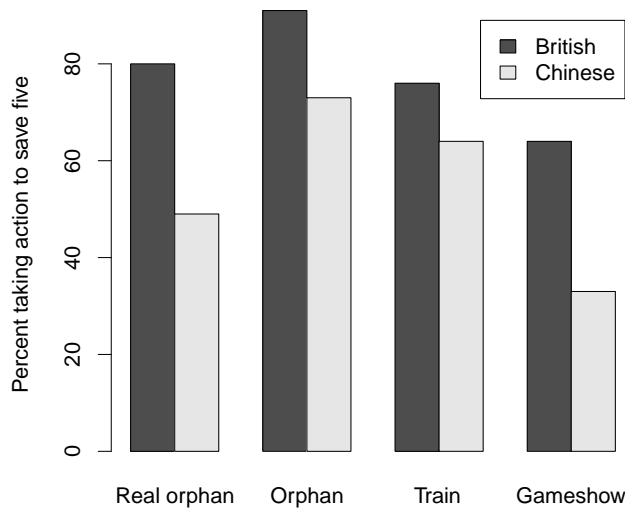
(3) I believe in fate

3.2 Results

3.2.1 Predicted decisions

In the Orphan scenario 90.91% of our British participants said that they would click the switch to save the five children from losing their meals, compared to 73.33% of Chinese (see Figure 2 for percentages for all scenarios, from both experiments). As in our previous experiment, the mean age of our Chinese participants was higher than the British. When we take account of this by running a logistic regression including demographic variables, we found that only the dummy variable for being Chinese was significant, with Chinese participants being less likely to click the switch than British. The model is shown in Table 5.

In the Train scenario more British (76.36%) than Chinese (64.44%) participants said that they would pull the lever to save the lives of the five people. A logistic regression including demographic variables shows an effect of nationality, with Chinese participants being less likely to say that they would move the lever to divert the train, and a trend for older participants to be more likely to say they would move it (see Table 5).

The same pattern also showed up in the responses to the Gameshow scenario, where 63.64% of the British said that they would press the button to save five contestants from

Figure 2: Percentage of participants in Experiments 1 and 2 choosing to take the action to save the five.

Table 5: Logistic regressions of predicted behavior (would click/ move/ press = 1, would not click/ move/ press = 0) with nationality (dummy variable for Chinese), gender (male = 0, female = 1), and age.

	B	SE	Wald	p	Exp(B)
Orphan					
Chinese	−1.85	0.71	6.87	.009	0.16
Age	0.20	0.13	2.11	.146	1.22
Gender	−0.36	0.65	0.30	.584	0.70
Constant	−1.45	2.78	0.27	.602	0.23
Train					
Chinese	−1.05	0.54	3.82	.051	0.35
Age	0.18	0.10	2.86	.091	1.19
Gender	0.23	0.50	0.21	.645	1.26
Constant	−2.65	2.20	1.45	.228	0.07
Gameshow					
Chinese	−1.26	0.48	6.94	.008	0.28
Age	0.00	0.09	0.00	.969	1.00
Gender	0.19	0.47	0.17	.680	1.21
Constant	0.35	1.86	0.04	.851	1.42

losing their money, but only 33.33% of the Chinese would do this. A logistic regression including demographic variables confirms that Chinese participants are less likely to press the button than British in the Gameshow scenario (see Table 5).

A more powerful test confirms our finding that the British were more likely to predict that they would take action than the Chinese. We combined the predictions for the three scenarios into a composite measure. We scored each response, as 0 for a "No" and 1 for a "Yes", and added each participant's three scores together, giving a measure that ranges from 0–3, α = .378. On our composite measure, British scores (M = 2.31) are higher than Chinese (M = 1.71), $t(74.9)$ = 3.37, p = .001.

3.2.2 Wrong-right ratings

In the Orphan scenario, there was a significant difference in the mean wrong-right ratings between British (M = 4.89) and Chinese (M = 4.07), $t(98)$ = 2.60, p = .011. When we run a multiple regression analysis that includes demographic variables, thus removing extraneous variance, it reveals that Chinese participants gave lower ratings, i.e., rated clicking the switch to be less right (see Table 6). Unlike our first experiment, there was no effect of age, but there was an effect of gender, with females rating the action as less right.

In the Train scenario, there was no difference in the mean wrong-right ratings between British (M = 4.18) and Chinese (M = 4.31), $t(98)$ = 0.34, p = .736. Regressions reveal that there were no cultural differences in ratings, although there was a gender effect, with female participants giving lower ratings (see Table 6).

In the Gameshow scenario, there was no difference in

the mean wrong-right ratings between British (M = 4.58) and Chinese (M = 4.22), $t(98)$ = 1.06, p = .292. Regressions show that there were no cultural differences or demographic effects on the numerical wrong-right judgment (see Table 6).

Analyzing a composite measure, which combines the responses from all three scenarios, shows an effect of nationality. Since the results from the three scenarios were in the same direction (even though some were n.s.), we created a composite measure, by adding together each participant's three wrong-right ratings. The measure ranges from 3–21, α = .632. The mean composite ratings for Chinese (M = 12.6) and British (M = 13.7) do not vary, $t(98)$ = 1.33, p = .186. However, when we remove extraneous variance by including the demographic variables in a regression, the effects of nationality and gender are both significant (see Table 6): Chinese and females give lower ratings.

3.2.3 Moral judgments

There were no cultural differences on our forced choice question, asking whether or not the action is morally wrong. In the Orphan scenario, 30.91% of British and 35.56% of Chinese judged that it would be morally wrong to click the switch, $\chi^2(1, 100)$ = .242, p = .623. In the Train scenario, 45.45% of British say 'Yes' it was wrong

Table 6: Multiple regression of wrong-right ratings on nationality (dummy variable for Chinese), gender (male = 0, female = 1), and age.

	B	SE	β	t	p
Orphan					
Chinese	-0.10	0.34	-0.31	-2.97	.004
Age	0.06	0.06	0.10	0.96	.340
Gender	-1.26	0.33	-0.35	-3.83	.000
Constant	4.60	1.31		3.51	.001
Train					
Chinese	-0.20	0.41	-0.05	-0.48	.631
Age	0.12	0.07	0.17	1.60	.113
Gender	-1.37	0.40	-0.33	-3.43	.001
Constant	2.74	1.59		1.73	.087
Gameshow					
Chinese	-0.57	0.38	-0.17	-1.48	.142
Age	0.08	0.07	0.13	1.14	.257
Gender	-0.25	0.38	-0.07	-0.65	.518
Constant	3.11	1.50		2.08	.041
Composite					
Chinese	-1.76	0.84	-0.22	-2.10	.039
Age	0.25	0.15	0.18	1.68	.097
Gender	-2.87	0.82	-0.33	-3.49	.001
Constant	10.5	3.28		3.19	.020

Table 7: Correlations of fate variables with predicted behavior (would click/ move/ press = 1, would not click/ move/ press = 0, summed), wrong−right ratings, and moral judgments (no it would not be morally wrong = 0, yes it would be morally wrong = 1, summed). N=99 (97 for Fate3 and composite); $p = .05$ for $|r| = .20$, .01 for $|r| = .26$, .001 for $|r| = .36$, two tailed.

	Predicted behavior	Wrong−right rating	Moral judgment
Composite fate	-0.32	-0.38	0.25
Fate1	-0.29	-0.26	0.17
Fate2	-0.35	-0.22	0.20
Fate3	-0.18	-0.39	0.27

posite measures reflect a similar pattern of significant correlations at the level of the individual scenarios (analysis available on request).

3.2.4 LOC and fate

The Chinese participants agreed much more strongly than the British participants that things should "run their natural course" (4.82 vs. 3.83, $t(97) = 2.80$, $p = .006$) [fate1], that it is "best not to try to interfere with the natural course of events", (4.49 vs. 3.37, $t(97) = 2.98$, $p = .004$) [fate2], and that they "believe in fate" (4.66 vs. 3.72, $t(95) = 2.32$, $p = .022$) [fate3]. These significant differences all show that the Chinese on average tend to believe in fate and that it is best to let events run their course and not intervene (higher than the mid-point of four on the scale) while the British do not tend to believe in fate and believe that they should intervene and not just let things happen. There were no significant differences in the Locus of Control scores of British (71.45) and Chinese (70.93) participants, $t(97) = 0.26$, $p = .796$.

Responses to the three fate questions were highly correlated with each other. A composite variable, which is the summation of each participant's three fate question ratings, ranges from 3–21, $\alpha = .785$. The three fate measures and the more reliable composite measure are all correlated with moral judgments and actions (Table 7). Belief in fate and that one should not take action is associated with a lower propensity to predict taking action, lower wrong-right ratings and a higher score on "is it morally wrong?"

3.2.5 Reasons

In all three scenarios, participants who said that they would take the action overwhelmingly cited utilitarian reasons about the greatest good of the greatest number, such as "for the greater good", "the lesser evil", "the needs of

vs. 35.56% of Chinese, $\chi^2(1, 100) = 1.00$, $p = .317$. In the Game scenario, 14.55% of British said that pushing the button would be morally wrong vs. 15.56% of Chinese, $\chi^2(1, 100) = .020$, $p = .888$.

This pattern is confirmed by a more powerful analysis of a composite measure. We scored each response, as 0 for a "No" and 1 for a "Yes", and added each participant's three scores together, making a composite measure that ranges from 0–3, $\alpha = .693$. On our aggregate measure, British scores ($M = .91$) are not significantly different from Chinese ($M = .87$), $t(98) = .20$, $p = .842$. Despite our Chinese sample being older than the British, there is a lack of significant correlation between age and moral judgment in the raw data and it is possible to confirm this lack of correlation with a logistic regression analysis (available on request).

However, the moral judgments are highly correlated with both the predicted action $r(100) = -0.326$, $p = .001$, and the wrong-right rating, $r(100) = -0.514$, $p < .001$. Participants who judged taking action to be morally wrong were less likely to predict that they would act and rated the action as more right. These correlations of the com-

the many outweigh the needs of the few", "to give more children more food" or "to save more children". This is true of both British and Chinese participants.

The reasons given by British and Chinese for not taking the action differed. In the Orphan scenario five British people said they would not click the switch. Three cited unclassifiable reasons, such as feeling sorry for the one child, or reasons that showed they had misunderstood the task. Two felt that the whole experimental set-up would be wrong and that therefore it would be wrong for them to choose saying, for example, "I think to even have such a scenario is wrong and would not wish to decide who should get a meal." In contrast, of the twelve Chinese participants who said that they would not click the switch, four cited reasons about not having the right to make this decision, such as "I have not the right to make this kind of decision", four cited reasons that we might classify as a "Kantian" equal respect for persons, such as "every life is equal to others" and "there is no evidence that 5 kids are more valuable than 1", one person said "I don't want to change or control anyone's destiny", and three were not possible to classify.

In the Train scenario, thirteen British participants said that they would not move the lever. The modal British reason for this, given by six participants, mentioned responsibility, such as "Did not want the responsibility of someones [sic] death." Of the sixteen Chinese participants who would not move the lever, the modal response, which was given by five people, cited destiny, fate, or nature; three of these also cited reasons regarding equal value of lives, such as "If it happens, it is destiny. It is not evident to say 5 worth more than 1." Four cited the fact that moving the lever would result in a death or a murder, and three said that they had no right to decide (of whom one also mentioned the equal value of lives).

In the Gameshow scenario, multiple reasons were generated. Of the twenty British who would not intervene, nine said that there were "no serious consequences of pressing or not pressing the button"; some of these mentioned that it was only £100 and others that it was not a life or death situation. Three said that they did not want to take part in the gameshow, two mentioned fate, and one said that that s/he didn't have any right to take the decision. In contrast, of the twenty-nine Chinese who said they would not press the button, six gave reasons that referenced fate or destiny or its not being their business to intervene; six said that they did not have the right to intervene, of whom three also mentioned fairness; six said that there were no serious consequences, six mentioned that it was only a gameshow; and two mentioned the harm to the one person who would lose his winnings.

4 Discussion

In our real-life trolley problem (Experiment 1 with orphans), the behavior of our British participants mirrored that of participants in an American virtual reality experiment (Navarette, McDonald, Mott, & Asher, 2012), and their moral judgments were in line with those made by the (mainly British) participants in hypothetical scenarios involving trolley problems associated with economic rather than mortal harms (Gold, Pulford, & Colman, 2013). However, a much smaller proportion of Chinese than British participants switched the lever, and fewer Chinese participants judged it to be right to switch the lever. Chinese participants were much more likely than British participants to judge this action neutrally, and this was in fact the modal Chinese judgment. The difference in propensity to take action and in wrong-right ratings was replicated in a second experiment using hypothetical scenarios, which British and Chinese completed in their own time and in their own languages. However, in there was no difference in dichotomous moral judgments of whether or not the action was morally wrong.

A higher percentage of participants of both nationalities said that they would take action in Experiment 2 than actually took action in Experiment 1. There was an increase amongst the British groups, from 80% in the real Orphan scenario to 91% in the hypothetical Orphan scenario, and an increase from 49% to 73% amongst the Chinese. Even if someone predicts that they would sacrifice one in order to save five, actually implementing that judgment may be more unpleasant and difficult than imagined. Alternatively, the difference might be due to timing constraints, as in Experiment 1 not clicking the switch within 11 seconds was considered to be a 'No', or due to differences between real and hypothetical decision-making. The larger increase amongst the Chinese group may also be because of better comprehension of the dilemma when it was in their own language. However, the difference between Chinese and British participants in taking action to save the five is still large and significant in both real and hypothetical situations.

In our second experiment, we also used a standard trolley scenario where lives were at stake, and our results confirm those of Moore, Lee, Clark and Conway (2011) and Mikhail (2011), as the Chinese and British participants' dichotomous moral judgments did not differ in the Train scenario. However, fewer Chinese than British participants say they are willing to actually take the required action. Across all three scenarios, the Chinese are less likely to say that they would take action and they rated taking action as less right than the British, but the two groups' dichotomous judgments did not differ. Given this pattern, it is not surprising that previous studies did not find cultural

differences because they only elicited moral judgments, whilst the main differences occur at the level of behavior.

Although there were no cultural differences in dichotomous moral judgments, judgments were strongly correlated with wrong-right ratings and with behavioural predictions. Both nationalities, on average, indicated that taking the action would be right, so the cultural difference represented a small shift along the "right" side of the ratings scale. Since the difference does not represent a shift from "right" to "wrong", this may explain why there was no cultural difference on the dichotomous measure of moral judgment.

Differences in demographics between our British and Chinese samples do not provide a good explanation of our results. Our participants were mainly students at a UK university. As well as having a different nationality, the Chinese group differed from the British in having chosen to study abroad and in being non-native English speakers (although they would have had to satisfy a test of proficiency in spoken and written English as a condition for admission). However, we controlled for the demographic confounds in our first experiment by using a group of other foreign students who were also not native English speakers, and which had very similar demographics to the Chinese group. The slightly higher mean age of both the Chinese and the other foreign group compared to the British may suggest a larger proportion of post-graduates. The other foreign group can control for the educational status of the Chinese, and age will also act as a proxy for status because number of years of education is correlated with age. Our control group participants did not make different decisions or judgments from the British, but they did differ from the Chinese. We also replicated our results regarding cultural differences in our second experiment where the materials were written in Chinese.

Cultural differences in fatalism provide a possible explanation of our results. Embedded in the "Great Tradition" of Chinese Taoism is a shared belief in fate (*ming* or *t'ien-ming*), interpreted as a force beyond human control that is chiefly responsible for determining people's destinies (Eberhard, 1966), and an associated ethical principle of action through non-action (*wu-wei*), or allowing events to take their natural course (Kirkland, 2004). Chinese fatalism has roots that can be traced back at least as early as the 8th century BC, and recent empirical studies have confirmed that it persists in contemporary Chinese societies, often in association with superstitious beliefs about numbers and colors, not only in mainland China, but also in Taiwan, Hong Kong, and other overseas territories (Chan, 2000), and even in Chinese communities in California (Phillips et al., 2001). Our Chinese participants were more likely to believe in fate, and that they should not intervene in the natural course of events. This is reflected in the ac-

tual reasons they gave, which tended to mention fate and destiny, and in the correlation of the answers to our fate questions with predicted actions, wrong-right ratings, and moral judgments. Further work would be needed to confirm this: a limitation of our design is that we asked several questions; the question about belief in fate came after the questions about non-intervention, which quite closely reflect the behavioural prediction that participants had already made, so the answer to the belief in fate question may have been contaminated by the prior questions and tasks.

Another possible explanation is that the Chinese and British differed in whether they thought they were responsible for taking action. Societies prescribe that certain decisions are to be made by particular people, who we might say have "responsibility" for the decision (Baron, 1996).

With respect to trolley problems, Thomson (1985) claimed that the driver would be in a special position of responsibility compared to a passenger; and people's moral judgments in trolley problems are correlated with their judgments about whether the agent is responsible for taking action (Gold, Pulford, and Colman, 2014b). Amongst the reasons that our participants gave for saying that they would not act, "not having the right to intervene" was a prevalent response from the Chinese. Their reluctance to act may be exacerbated by the fact that Chinese have more inter-dependent self-construals, one consequence of which is that they care more about the opinion of others (Markus & Kitayama, 1991). The Chinese may have been more worried about being negatively perceived by others if they caused harm to someone when taking a decision that they felt they had no right to make.

The difference in behavior between our British and Chinese samples is clear enough. How the behavioral difference relates to the difference in moral judgments, on the other hand, must be interpreted with care.

In our real life Orphan scenario, moral judgments generally corresponded with actions, and we may indeed have elicited the moral judgments that underpinned the actions. However, it is possible that participants were motivated to report judgments consistent with their actions. This could result from conscious misreporting, motivated by social desirability and image management and intended to convey an impression of consistency. Alternatively, participants may have reported their judgments truthfully but, because judgments were elicited after actions, they may have tended to form judgments that were consistent with actions previously taken in order to avoid cognitive dissonance (Brehm, 1956; Gawronski & Strack, 2004; Stone & Cooper, 2001). However, the results of Experiment 2 speak against a cognitive dissonance explanation. If dissonance were an issue, then we would expect that a decision to act would be followed by a judgment that acting

was not morally wrong. Instead, in the hypothetical Orphan and the Train scenarios, strikingly more participants of both nationalities judged that it would be morally wrong to act than predicted that they would refrain from acting.

Our findings also raise issues about the consistency between moral judgments and moral behavior in trolley problems. Tassy et al. (2013) also found that people's predictions of their actions were more utilitarian than their normative judgments. They hypothesize that the difference is caused because judgments and choices are result from (at least partially) different psychological processes. We also suspect that our participants did not see moral considerations as over-riding reasons for action. Rather, they were only one consideration that could be outweighed. But further investigation is needed to say anything definitive about the causes of the difference between moral judgment and behavior, or the cultural differences in moral behaviour.

References

Abarbanell, L., & Hauser, M. D. (2010). Mayan morality: An exploration of permissible harms. *Cognition, 115,* 207–224.

Ahlenius, H., & Tännsjö, T. (2012). Chinese and Westerners respond differently to the trolley dilemmas. *Journal of Cognition and Culture*, *12*, 195–201.

Baron, J. (1996). Do no harm. In D. M. Messick & A. E. Tenbrunsel (Eds.), *Codes of conduct: Behavioral research into business ethics*, pp. 197–213. New York: Russell Sage Foundation.

Bartels, D. M. (2008). Principled moral sentiment and the flexibility of moral judgment and decision making. *Cognition, 10*, 381–417.

Bond, M. H. (1991). *Beyond the Chinese face: Insights from psychology.* Hong Kong: Oxford University Press.

Brehm, J. (1956). Post-decision changes in the desirability of alternatives. *Journal of Abnormal and Social Psychology, 52,* 384–389.

Chan, W. S. (2000). *Chinese fatalism and its relation to coping and adaptation outcomes* (Master's thesis, University of Hong Kong). Retrieved from http://hub.hku.hk/handle/10722/33311

Chuangsheng, C., Shin-Ying, L., & Stevenson, H. (1995). Response style and cross-cultural comparisons among rating scales of east Asian and north American students. *Psychological Science, 6*, 170–175.

Cushman, F., Young, L., & Hauser, M. D. (2006). The role of conscious reasoning and intuition in moral judgment: Testing three principles of harm. *Psychological Science*, *17*, 1082–1089.

Eberhard, W. (1966). Fatalism in the life of the common man in non-communist China. *Anthropological Quarterly, 39,* 148–160.

Gawronski, B., & Strack, F. (2004). On the propositional nature of cognitive consistency: Dissonance changes explicit, but not implicit attitudes. *Journal of Experimental Social Psychology, 40,* 535–542.

Gold, N., Pulford, B. D., & Colman, A. M. (2013). Your money or your life: Comparing judgements in trolley problems involving economic and emotional harms, injury and death. *Economics and Philosophy, 29,* 213–33.

Gold, N., Pulford, B. D., & Colman, A. M. (2014a). Do as I say, don't do as I do: Differences in moral judgments do not translate into differences in decisions in real-life trolley problems. Manuscript submitted for publication

Gold, N., Pulford, B. D., & Colman, A. M. (2014b). The outlandish, the realistic, and the real: Contextual manipulation and agent role effects in trolley problems. *Frontiers in Psychology: Cognitive Science, 5*, 35.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe*, Vol. 7 (pp. 7–28). Tilburg, The Netherlands: Tilburg University Press.

Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, *111*, 364–71.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, *44*, 389–400.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293,* 2105–2108.

Hare, R. M. (1952). *The language of morals*. Oxford: Oxford University Press.

Hauser, M. D., Cushman, F. A., Young, L., Jin, R. K.-X., & Mikhail, J. (2007). A dissociation between moral judgments and justifications. *Mind & Language*, *22*, 1–21.

Henrich, J., Heine, S. J., Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33,* 61–83.

Hsu, M., Anen, C., & Quartz, S. R. (2008). The right and the good: Distributive justice and neural encoding of equity and efficiency. *Science, 320,* 1092–1095.

Kamm, F. M. (2007). *Intricate Ethics*. New York: Oxford University Press.

Kirkland, R. (2004). *Taoism: The enduring tradition.* London: Routledge.

Machery, E., Mallon, R., Nichols, S., & Stich, S. P. (2004). Semantics, cross-cultural style. *Cognition*, *92*, B1-B12.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98,* 224-253.

Mikhail, J. (2011). *Elements of moral cognition: Rawls' linguistic analogy and the cognitive science of moral and legal judgment.* New York: Cambridge University Press.

Moore, A., Clark, B., & Kane, M. (2008). Who shalt not kill? Individual differences in working memory capacity, executive control, and moral judgment. *Psychological Science, 19,* 549–557.

Moore, A. B., Lee, N. Y. L., Clark, B. A. M., & Conway, A. R. A. (2011). In defense of the personal / impersonal distinction in moral psychology research: Cross-cultural validation of the dual process model of moral judgment. *Judgment and Decision Making*, *6,* 186–195.

Navarrete, C. D., McDonald, M., Mott, M., & Asher, B. (2012). Virtual morality: Emotion and action in a simulated 3-D "trolley problem." *Emotion, 12*, 364–370.

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic vs. analytic cognition. *Psychological Review, 108*, 291-310.

O'Neill, P., & Petrinovich, L. (1998). A preliminary cross-cultural study of moral intuitions. *Evolution and Human Behavior, 19*, 349–367.

Osberg, T. M., & Shrauger, J. S. (1986). Self-prediction: Exploring the parameters of accuracy. *Journal of Personality and Social Psychology, 51,* 1044–1057.

Palsane, M. N., & Lam, D. J. (1996). Research on stress and coping: Contemporary Asian approaches. In A. C. Paranjpe, D. Y. F. Ho, & R. W. Rieber (Eds.), *Asian contributions to psychology* (pp. 265–281). New York: Praeger.

Phillips, D. P., Liu, G. C., Kwok, K., Jarvinen, J. R., Zhang, W., & Abramson, I. S. (2001). The *Hound of the Baskervilles* effect: Natural experiment on the influence of psychological stress on timing of death. *British Medical Journal, 323,* 22–29.

Schaich Borg, J., Hynes, C., Van Horn, J., Grafton, S., & Sinnott-Armstrong, W. (2006). Consequences, action, and intention as factors in moral judgments: An fMRI investigation. *Journal of Cognitive Neuroscience, 18,* 803–817.

Schuman, H., & Presser, S. (1981). *Questions and answers in attitude surveys.* New York: Academic Press.

Schwitzgebel, E., & Rust, J. (2013). The moral behavior of ethics professors: Relationships among self-reported behavior, expressed normative attitude, and directly observed behavior. *Philosophical Psychology*, (ahead-of-print), 1–35.

Stone, J., & Cooper, J. (2001). A self-standards model of cognitive dissonance. *Journal of Experimental Social Psychology, 37,* 228–243.

Tassy, S., Oullier, O., Mancini, J., & Wicker, B. (2013). Discrepancies between judgment and choice of action in moral dilemmas. *Frontiers in Psychology, 4*, No. 250. http://dx.doi.org/10.3389/fpsyg.2013.00250.

Teper, R., Inzlicht, M., & Page-Gould, E. (2011). Are we more moral than we think?: Exploring the role of affect in moral behavior and moral forecasting. *Psychological Science, 22*, 553-558.

Thomson, J. J. (1985). The trolley problem. *Yale Law Journal, 94,* 1395–1415.

Vallone, R. P., Griffin, D .W., Lin, S., & Ross, L. (1990). Overconfident prediction of future actions and outcomes by self and others. *Journal of Personality and Social Psychology, 58,* 582–592.

Waldmann, M. R., & Dieterich, J. H. (2007). Throwing a bomb on a person versus throwing a person on a bomb: Intervention myopia in moral intuitions. *Psychological Science, 18,* 247–253.

Weber, E. U., & Morris, M. W. (2010). Culture and judgment and decision making: The constructivist turn. *Perspectives on Psychological Science*, *5*, 410–419.

Weinberg, J., Nichols, S., & Stich, S. (2001). Normativity and epistemic intuitions. *Philosophical Topics, 29*, 429–459.