

SPECIAL ISSUE ARTICLE

Effects of dilemma type, language, and emotion arousal on utilitarian vs deontological choice to moral dilemmas in Chinese–English bilingualsYuen-Lai Chan,¹ Xuan Gu,¹ Jacky Chi-Kit Ng² and Chi-Shing Tse¹*Departments of ¹Educational Psychology and ²Psychology, The Chinese University of Hong Kong, Hong Kong*

The present study examines how dilemma type (personal or impersonal moral dilemma), language (native or foreign) and emotion arousal to a dilemma could affect Chinese–English bilinguals' deontological vs utilitarian moral choices regarding 39 moral dilemmas. How emotion arousal plays a mediating role in the effects of dilemma type and language on moral choices is also investigated. As shown in multilevel analyses, participants made fewer utilitarian choices for personal dilemmas than impersonal dilemmas. Although emotion arousal of dilemmas significantly mediated this effect of dilemma type, the indirect effect of dilemma type through arousal on moral choices was inconsistent with the direct effect of dilemma type on moral choices. For the effect of language, participants made more utilitarian choices in the Footbridge (personal) dilemma that was presented in foreign language than in native language. However, this effect was not mediated by arousal, suggesting that it could not be attributed to the emotion-reducing effect of foreign language. Moreover, there was no language effect on moral choices in analyses that included all 39 dilemmas or only 22 personal dilemmas, indicating the need in future research for further identifying the potential mediators that trigger the foreign language effect on moral choices.

Key words: deontology, emotion, language, moral dilemma, utilitarian.

Several factors have been identified that influence people's decisions regarding moral dilemmas, such as dilemma type (i.e. personal or impersonal; e.g. Greene, Sommerville, Nystrom, Darley & Cohen, 2001; Moore, Lee, Clark & Conway, 2011), language in which the dilemmas are presented (e.g. Cicolletti, McFarlane & Weissglass, 2015; Costa *et al.*, 2014; Geipel, Hadjichristidis & Surian, 2015) and emotional characteristics of the dilemmas (e.g. Carmona-Perera, Martí-García, Pérez-García & Verdejo-García, 2013; Pizarro, 2000; Ugazio, Lamm & Singer, 2012). Most previous findings, however, are based on participants in Western societies and dilemmas presented in Western languages (e.g. English, Spanish and German), and no published study has investigated how these factors jointly affect judgement regarding moral dilemmas. Previous studies often considered emotion arousal as playing a mediating role in the effects of dilemma type and language on moral choices. For example, people may feel more distant (and thus less emotional) when reading personal dilemmas that are presented in a foreign language than those that are presented in a native

language (e.g. Cicolletti *et al.*, 2015; Costa *et al.*, 2014). Nevertheless, emotion arousal has rarely been directly measured, and its mediating role has not been explicitly tested in the literature. On the other hand, previous studies that reported the foreign language effect on moral choices often attributed that to the emotion-reducing effect of language. However, an alternative account (cultural-priming effect of language) should also be considered. To address these concerns and further test the theoretical accounts for the processing of moral dilemmas, in the current study we asked Chinese–English bilinguals in Hong Kong to rate their emotion arousal triggered by moral dilemmas that were presented in either their native language (Chinese) or a foreign language (English) and to make moral judgements between utilitarian vs deontological choices.¹

The present study has three goals. First, we aim to explore how dilemma type, language and emotion arousal of moral dilemmas would separately and/or jointly influence bilinguals' moral choices within one multilevel statistical model. Second, we aim to examine whether dual-process theory (e.g. Greene *et al.*, 2001), a widely accepted theory in moral judgement, could explain the effect of dilemma type on bilingual participants' moral choices in Hong Kong. Third, we aim to test the language effect on moral choices and verify whether emotion arousal plays a mediating role in the effect of language on moral choices. Before reporting the details of the current study, we first

Correspondence: Yuen-Lai Chan, Department of Educational Psychology, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong. Email: heidichan@cuhk.edu.hk
Received 28 February 2015; revision 27 October 2015; accepted 16 November 2015.

introduce the dual-process theory of moral judgement to illustrate the relationship among dilemma type, emotion arousal and moral choices. We then introduce and distinguish two potential effects of language on moral choices, that is, the emotion-reducing effect of foreign language *vs* the culture priming effect of language.

The dual-process theory of moral judgement

According to Greene and his colleagues (2008, 2004, 2001) the dual-process theory of moral judgement, utilitarian choices, which approve of potentially harmful actions to maximize good consequences, are driven by more cognitive, controlled processes, whereas deontological choices, which disapprove of any potentially harmful actions on affected individuals, are driven by emotional, automatic responses. The respective roles of emotional *vs* cognitive processes in moral judgements are often examined in laboratories via a moral dilemma paradigm (e.g. Greene, Morelli, Lowenberg, Nystrom & Cohen, 2008; Greene, Nystrom, Engell, Darley & Cohen, 2004; Greene *et al.*, 2001). In this paradigm, each dilemma involves a scenario and two possible resolutions, which reflect either a utilitarian choice that maximizes benefits and minimizes costs or a deontological choice that conveys emotional responses. Participants read the dilemmas as if they were the protagonist in the scenario and choose either resolution. There are two types of moral dilemma. In personal dilemmas, moral violations should meet three criteria: (i) they cause serious bodily harm; (ii) the harm occurs to a particular person or group; and (iii) the harm is not a result of the deflection of an existing threat (e.g. Greene, 2009; Greene *et al.*, 2004). Dilemmas that do not meet any of these criteria are classified as impersonal.

Take the Trolley dilemma (Foot, 1967) and Footbridge dilemma (Thomson, 1985) as examples, which are the most representative impersonal and personal dilemmas, respectively. In the Trolley dilemma you are inside the control room of a trolley that is approaching five people on one side, and there is one person on the other side. You need to choose between (i) pressing a button to make the trolley turn to the other side to kill one person and save five people and (ii) not pressing the button, so that the person on the other side survives but the five people will die. In the Footbridge dilemma you are standing on a bridge and you see a trolley is approaching five people. You need to choose between (i) pushing a big stranger next to you to stop the trolley, so that the stranger will die but the five people will be saved and (ii) doing nothing, so that the stranger will survive but the five people will die. The Footbridge dilemma is personal because it includes serious harm on a particular person's body, whereas the Trolley dilemma is impersonal as it involves the deflection of an existing threat.

In both dilemmas, the first judgement is to make a utilitarian choice to sacrifice one person in order to save five and the second judgement is to make a deontological choice, which reasons that it is immoral to kill any innocent person.

There has been evidence supporting the dual-process theory in terms of functional magnetic resonance imaging (fMRI) and reaction time (RT) data (e.g. Greene *et al.*, 2001, 2004, 2008; Koenigs *et al.*, 2007; Moore, Clark & Kane, 2008; Moore *et al.*, 2011; Young *et al.*, 2010; Young & Koenigs, 2007). For example, emotion-related brain regions (e.g. ventromedial prefrontal cortex) increased activation when participants made deontological choices (Greene *et al.*, 2001). In contrast, when utilitarian choices were made, brain regions associated with cognitive control (e.g. dorsolateral prefrontal cortex) were activated (Greene *et al.*, 2004). Different patterns of brain activity associated with deontological *vs* utilitarian choices suggest that deontological choices were likely driven by emotional processes, whereas utilitarian choices are likely to be driven by cognitive processes. Moretto, Ládavas, Mattioli and Di Pellegrino (2010) used skin conductance to measure participants' emotional responses when making moral judgements. Specifically, they compared the skin conductance responses of patients with damaged emotion-related brain regions with those of control participants. In contrast to the control participants who made more deontological choices in personal dilemmas, the patients with damaged emotion-related brain regions made more utilitarian choices in personal dilemmas and failed to generate any emotional responses when making the choices. This again indicates the involvement of emotional processing when making deontological choices in personal dilemmas.

For the effects of emotion arousal and valence on moral judgements, Christensen, Flexas, Calabrese, Gut and Gomila (2014) examined the arousal and valence of 46 moral dilemmas presented in Spanish and found that dilemmas with higher arousal ratings yielded more deontological choices. (There was no effect of the dilemmas' valence on moral judgements.) Lotto, Manfrinati and Sarlo (2014) report normative data for 72 dilemmas presented in Italian, including moral acceptability of utilitarian choice, proportion of utilitarian choice, participants' RTs, and arousal and valence ratings for the moral judgements. Instead of using the personal *vs* impersonal distinction, Lotto *et al.* categorize the dilemmas as incidental (which involved an action with unintended harm) or instrumental (which involved an action with intended harm). They found that utilitarian choices were rated as more permissible for the incidental dilemmas than for the instrumental dilemmas, and judgements to incidental dilemmas were rated to be *more* emotionally arousing than those to instrumental dilemmas. (The two types of dilemma were comparable in valence rating.) However, Lotto *et al.*'s participants were instructed to rate the arousal and valence for their moral choices rather

than the immediate emotions that they felt towards the dilemmas.

Previous research (e.g. Greene *et al.*, 2004) argues that personal dilemmas could trigger emotional processing, such that participants are more likely to make deontological choices regarding personal dilemmas. However, they did not directly test whether the emotion arousal of the moral dilemmas *per se* could influence participants' moral choices. That is, it is not clear whether the more arousing dilemmas would lead to more deontological choices than the less arousing ones. To test the influence of emotion arousal elicited by the dilemmas on utilitarian and deontological moral choices, in the current study we had participants rate the arousal (i.e. emotional intensity) and valence (i.e. pleasantness) based on the moral dilemmas before they made a choice. That is, immediately after they read the dilemmas, participants rated the emotion arousal and valence triggered by the dilemmas. Unlike previous studies that did not measure participants' ratings of emotion arousal triggered by moral dilemmas (e.g. Greene *et al.*, 2004) or only measured participants' ratings of emotion arousal triggered by their moral choices (e.g. Lotto *et al.*, 2014), the design of the current study allowed us to measure the emotion triggered by the dilemmas and how it affected moral choices. While controlling for the valence rating of dilemmas, we also tested whether arousal rating would mediate the relationship between dilemma type (personal vs impersonal) and participants' moral choices. This could directly lead to a hypothesis derived from the dual-process theory, that is, deontological choices are driven by emotional processing of more arousing, personal dilemmas, whereas utilitarian choices are driven by cognitive processing of less arousing, impersonal dilemmas.

The 'language effect' on moral judgement

The effect of language on moral judgement has recently drawn increasing attention in the literature (see Caldwell-Harris, 2015, for a review; e.g. Cicolletti *et al.*, 2015; Costa *et al.*, 2014; Geipel *et al.*, 2015). There could be two possible effects of language on moral choice, although they have rarely been teased apart in previous studies. First, language may affect moral judgement because foreign language, compared to native language, is more likely to be associated with lower emotion arousal. Second, language used in the moral dilemma task could influence moral judgement via cultural priming.

The emotion-reducing effect of foreign language. In Keysar, Hayakawa and An (2012), one of the first studies to report the influence of foreign language on decision making, bilingual participants made judgements in their native or foreign language in various decision-making tasks that involved framing risks or loss aversion. The results

show that responding in a foreign language reduced participants' decision-making biases (e.g. reduced loss aversion). Keysar *et al.* explains that a foreign language may provide a greater distance than a native language and thus lead participants to trigger more cognitive processing than emotional processing in decision making. However, it is unclear whether Keysar *et al.*'s foreign language effect could be attributed to a reduced emotion arousal when the betting games were presented in a foreign language because this variable was not directly measured in their study.

Investigating the foreign language effect on specifically moral judgements, Costa *et al.* (2014) tested whether bilinguals made different judgements when moral dilemmas were presented in their native language (Korean, English or Spanish) or a foreign language (English, Spanish, Hebrew or French). The participants made fewer deontological choices in the Footbridge dilemma (a personal dilemma) when it was presented in a foreign language than in a native language. However, this foreign language effect did not occur in the Trolley dilemma (an impersonal dilemma). Cicolletti *et al.* (2015) replicated this effect when limiting the native–foreign language combination to English and Spanish.

Both Costa *et al.* (2014) and Cicolletti *et al.* (2015) attribute the foreign language effect on moral judgements to a reduction of emotion arousal. Because emotion arousal is *presumably* higher in personal dilemmas and foreign language might reduce automatic emotional responses associated with personal dilemmas, participants were more likely to make utilitarian choices for personal dilemmas presented in a foreign language than those presented in a native language. This effect did not occur in impersonal dilemmas because they were less aroused than personal dilemmas. Nevertheless, this possibility was not tested in either study because they did not measure the emotion arousal of their moral dilemmas. In the current study, we asked Chinese–English bilingual participants to rate their immediate emotions elicited by the moral dilemmas presented in native or foreign languages. In addition, unlike in previous studies, a larger set of moral dilemmas was used to examine whether the foreign language effect could generalize to other personal and impersonal dilemmas. If Costa *et al.* and Cicolletti *et al.*'s findings could have been replicated in a larger set of moral dilemmas, participants would have made more utilitarian choices in personal dilemmas when they were presented in English than in Chinese. If the foreign language effect can be attributed to the reduction of arousal in personal dilemmas, the aforementioned relationship between foreign language and utilitarian choices would be mediated by the arousal of personal dilemmas. In contrast, participants' judgements regarding impersonal dilemmas would not be affected by the language of the dilemmas. However, apart from the emotion-reducing effect of foreign language, it is important to point out an

alternative, cultural-priming effect of language, which could also drive the effect of language on moral choices.

The culture priming effect of language. According to the Sapir-Whorf linguistic relativity hypothesis (Sapir, 1949; Whorf, 1956), culture influences people's thinking through language. The relationship among language, culture and cognition have long been discussed (e.g. Chiu, Leung & Kwan, 2007; Oyserman & Lee, 2007). Chinese-speaking societies are generally more collectivistic than English-speaking societies (e.g. Oyserman & Lee, 2008; Triandis, McCusker & Hui, 1990) with the former valuing group interest while the latter values personal interest. Hong, Morris, Chiu and Benet-Martinez (2000) suggest that the cultural values of bicultural individuals could be switched by priming methods, such as language priming. Oyserman and Lee's (2008) meta-analysis also suggests that collectivistic and individualistic cultures could be induced by language priming.

As Chinese-English bilinguals in Hong Kong are also bicultural, and are influenced by both Eastern (collectivistic) and Western (individualistic) cultures simultaneously, it is possible that their native (Chinese) and foreign (English) languages might prime different sets of cultural values, which would modulate the effect of native or foreign language on moral judgement. Earlier research did show that Hong Kong students with different language exposure were associated with different sets of cultural values (e.g. Earle, 1969). For decision making, Briley, Morris and Simonson (2005) found that Hong Kong students tended to avoid losses when the decision task was presented in Chinese (vs English), complying with the prevention focus strategy of the motivational approach in Chinese culture. For decision making in the realm of morality, Mann and Cheng (2013) examined the inter-relationship among cultural values, empathy and moral competence in Hong Kong bilingual students. They focused on participants' moral competence instead of their judgements of moral dilemmas and found that empathy interacted with the degree of collectivism in predicting moral competence. This suggests that a cultural factor like collectivism could modulate moral reasoning in bilinguals. These studies motivated us to hypothesize a potential cultural-priming effect of language for our Chinese-English (bicultural) bilingual participants.

Studying the relationship between self and culture, Markus and Kitayama (1991) suggest that people in Eastern and Western cultures have different self-construal (dependent or interdependent), which affects one's cognition and emotion. Regarding moral reasoning, two systems, duty-based and right-based morality, can be applied to explain how people differ in moral reasoning in different cultures (e.g. Chiu, Dweck, Tong & Fu, 1997; Sachdeva, Singh & Medin, 2011). According to Sachdeva *et al.*, people who believe in duty-based morality are more likely to consider

the group interest during moral reasoning, so they would make more utilitarian choices. On the contrary, people who believe in right-based morality are more likely to consider their own interest and thus make more deontological choices. The difference between Eastern Asian and North American in duty- vs right-based thinking may impact duty- vs right-based moral reasoning. Cross-cultural studies (e.g. Hong & Chiu, 2001; Hong, Ip, Chiu, Morris & Menon, 2001) demonstrate that Chinese people maintained duty-based morality while American people maintained right-based morality. Based on this evidence and the cultural difference in collectivism and individualism, we predict that if Chinese language primes collectivistic value/duty-based morality, participants will prefer common good over the rights of individuals when the dilemmas are presented in Chinese, such that the Chinese language would yield more utilitarian choices. This result would be observed in both personal and impersonal dilemmas. This hypothesis is in contrast to the one derived from the emotion-reducing effect of language. To reiterate, if foreign language reduces the emotion arousal of personal dilemmas, we expect that participants are more likely to make utilitarian choices in personal dilemmas (but not in impersonal dilemmas) when they are presented in English.

Overview of the present study

To recapitulate, there are three goals in the current study. First, we aim to explore how dilemma type, language and emotion arousal contribute to moral choice within a statistical model. Second, we aim to replicate previous findings on the dual-process theory of moral judgement in a bicultural community and examine whether Hong Kong participants would make more deontological choices in personal dilemmas, but more utilitarian choices in impersonal dilemmas. If this prediction is supported, we will test if the relationship can be mediated by the arousal ratings of the dilemmas. Third, we aim to explore the language effect on moral judgement and find out whether the effect is due to the emotion-reducing effect of foreign language or the cultural-priming effect for bicultural/bilingual participants in Hong Kong. On one hand, if the self-rated emotion arousal of moral dilemmas matters, bilingual participants would make more utilitarian choices when the personal dilemmas are presented in a foreign language (English). On the other hand, if collectivistic value/duty-based morality or individualistic value/right-based morality can be activated by language, bilingual participants would make more utilitarian choices when the dilemmas are presented in Chinese, regardless of (personal or impersonal) dilemma type.

Unlike previous studies (e.g. Christensen *et al.*, 2014; Costa *et al.*, 2014; Greene *et al.*, 2001), our participants rated the arousal and valence elicited by the personal and

impersonal dilemmas to shed light on the relationship between participants' emotional responses and their judgements regarding moral dilemmas. These emotion ratings, together with other questions/ratings of the dilemmas (e.g. participants' previous knowledge about the dilemma and imagery vividness), could serve as normative data for future studies on moral judgement in the Chinese community (see the supporting information). On the other hand, previous research in decision making shows that question type could induce a framing bias (Pastötter, Gleixner, Neuhauser & Bäuml, 2013). Pastötter *et al.* found that how emotion affected moral judgement depended on how the questions were presented to participants. There were two types of question in their study: one was an active frame, 'Is it appropriate to push?' and the other was a passive frame, 'Is it appropriate not to push?'. They found that participants in the positive emotion answered 'yes' in both active and passive frames; that is, they made utilitarian choices in both frames. To minimize any interaction effect between emotion and question type, we asked 'What would you do?' after participants read the dilemmas and provided two options that reflected the utilitarian and deontological choices, instead of providing yes/no options or the permission/wrongness scale. Furthermore, we included comprehension checks after participants made a choice in each dilemma in order to make sure that the dilemmas were well understood. Only responses from participants who answered the two comprehension-check questions correctly could reflect the moral judgement processes and were included in our data analyses.

We propose the following hypotheses with respect to the three goals:

- 1 To explore the effects of dilemma type, language and emotion arousal on moral choice.
 - 1a) Personal dilemmas would yield more deontological choices while impersonal dilemmas would yield more utilitarian choices (e.g. Greene *et al.*, 2008).
 - 1b) Dilemmas presented in Chinese (native language) would yield more deontological choices while those in English (foreign language) would yield more utilitarian choices (e.g. Costa *et al.*, 2014).
 - 1c) More arousing dilemmas would yield more deontological choices while less arousing dilemmas would yield more utilitarian choices (e.g. Christensen *et al.*, 2014).
- 2 To test the dual-process theory of moral judgement.
 - 2a) The dilemma-type effect on moral choice would be mediated by emotion arousal such that personal dilemmas would be rated higher in emotion arousal, and higher emotion arousal would yield fewer utilitarian choices (e.g. Greene *et al.*, 2008).
- 3 To examine whether the language effect on moral choice would be due to the emotion-reducing effect or the cultural-priming effect.

- 3a) If foreign language attenuates the emotions (as quantified by emotion arousal) in personal dilemmas, dilemmas presented in English would yield more utilitarian choices than deontological choices, relative to those presented in Chinese (e.g. Costa *et al.*, 2014).
- 3b) If native and foreign languages prime different cultural values for bilingual participants, more utilitarian choices would be made when the dilemmas are presented in Chinese (vs English) as Chinese would prime collectivistic values and interdependence.
- 3c) If the emotion-reducing effect of foreign language in personal dilemmas is found, the relationship between language and moral choice would be mediated by emotion arousal, such that dilemmas presented in English would be less arousing, and these lower arousing dilemmas would yield more utilitarian choices.

Methods

Participants and design

One hundred and forty-four (75 females) Chinese–English bilingual undergraduates from the Chinese University of Hong Kong participated in exchange of monetary compensation (~6.5 USD). Their ages ranged from 18 to 26 ($M = 20.50$, $SD = 1.62$). They were randomly assigned to Chinese (native language) and English (foreign language) groups, for which all stimuli were presented in Chinese and English, respectively. There was no significant difference in English language proficiency across the two groups, $p_s > 0.10$.

In addition to the two experimental groups, we recruited another two groups for validity purposes in our study, namely the validation groups. The validation groups were included because one may argue that our participants did not make moral choices immediately after reading the dilemmas, which would bypass the automatic process and therefore bias the judgement. By testing if the group status (experimental vs validation) would interact with any factors on participants' moral choices, we could verify whether the inclusion of emotion rating prior to making moral choices would significantly influence the results that we obtained from the experimental group. The procedure for the experimental and validation groups was the same, except that those in the experimental groups made moral choices *after* emotion ratings whereas those in the validation groups did not rate the emotion, such that they made moral choices *immediately* after reading the dilemmas. The validation groups consisted of 144 Chinese–English bilingual undergraduates (72 females, age $M = 19.69$, $SD = 1.56$) from the same population as the experimental groups. They partici-

Table 1 Participants' demographic information

	Sample size	Age <i>M (SD)</i>	English language proficiency			
			Age of acquisition <i>M (SD)</i>	Speaking† <i>M (SD)</i>	Understanding† <i>M (SD)</i>	Reading† <i>M (SD)</i>
<i>Experimental Groups</i>						
Chinese/native language	72	20.35 (1.42)	3.96 (2.09)	6.26 (1.57)	6.47 (1.52)	6.82 (1.55)
English/foreign language	72	20.65 (1.79)	4.07 (2.51)	6.18 (1.66)	6.68 (1.66)	6.83 (1.60)
<i>Validation Groups</i>						
Chinese/native language	72	19.49 (1.53)	3.61 (1.53)	6.35 (1.60)	6.28 (1.76)	6.81 (1.62)
English/foreign language	72	19.90 (1.57)	3.61 (1.64)	6.58 (1.42)	6.76 (1.61)	7.17 (1.53)

[†]Rating scales of proficiency on speaking, understanding and reading were ranged from 0 to 10.

pated in exchange for monetary compensation (~6.5 USD) and were randomly assigned to Chinese (native language) and English (foreign language) groups. They did not significantly differ in English language proficiency compared with the experimental groups, $p_s > 0.10$. Table 1 presents participants' demographic information for the experimental and validation groups.

Materials and procedure

The moral dilemma task was conducted via E-prime 2.0 (Schneider, Eschman & Zuccolotto, 2002). The 22 personal and 17 impersonal dilemmas were adopted from Greene and his colleagues (2008, 2004, 2001). The moral dilemmas were translated from English to Chinese by five Chinese–English bilingual students from the same population as our participants and checked by all authors (see supporting information). The following procedure was identical for all participants, except that those in the validation groups did not rate the arousal and valence triggered by the dilemmas. The dilemmas were presented randomly and a non-moral dilemma was given as a practice trial at the beginning of the study. Each dilemma consisted of a scenario and two resolutions that appeared on the screen simultaneously until participants pressed the ENTER key to proceed. Participants' reading time and moral choice for each moral dilemma were recorded.

Following the offset of a dilemma, participants were asked to answer the following questions related to the dilemma.

- 1 The emotion triggered by the dilemma was measured by the Self-Assessment Manikin scale (Bradley & Lang, 1994). Participants rated the emotion arousal and valence of the dilemma on a nine-point Likert scale (1 = least arousing/pleasant, 9 = most arousing/pleasant). The scale was presented on the screen and on an instruction sheet. The participants were instructed to rate the emotion based on their immediate experience towards the dilemma. The presentation order of the emotion arousal and valence ratings was randomized across dilemmas.

- 2 The moral choice question (what would you do?) consisted of two options that reflected utilitarian and deontological actions corresponding to the resolutions. Whether the utilitarian or deontological choice appeared above the other was counterbalanced between participants.
- 3 The comprehension checks of participants' understanding in each dilemma included two questions (each with two options) related to the dilemma. For example, in the Footbridge dilemma, participants were asked: (i) 'What would happen if you pushed the stranger?' and (ii) 'What would happen if you do not push the stranger?' The two options followed were 'The stranger will live' and 'The five workmen will live', and 'The stranger will die' and 'The five workmen will die', respectively. The dilemmas appeared simultaneously with the comprehension checks on the screen because we intended to test participants' understanding, rather than memory, of the dilemmas.
- 4 Participants' previous knowledge about the dilemmas was measured by asking whether they had known about each dilemma before (1 = yes, 0 = no).
- 5 The vividness of the image related to the dilemma was measured on a seven-point Likert scale, with '1 = not vivid at all' and '7 = extremely vivid'.

All questions were presented one at a time and stayed on the screen until participants responded. Participants were allowed to take a short self-pace break after responding to each dilemma. At the end of the whole task, participants were asked to report the age of acquisition for English and rate their levels of proficiency in speaking, understanding and reading, on a ten-point Likert scale (0 = none, 10 = perfect), as compared with a native English speaker (see Table 1).

Results

The following analyses were based on the data of the experimental groups, who performed the emotion ratings

Table 2 Mean† and Standard Deviation† of the emotion ratings, reading time and the imagery vividness of the dilemmas (experimental groups)

		Arousal‡		Valence‡		Reading Time (ms)		Imagery Vividness§	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Chinese/native language	Personal dilemmas ($N = 22$)	5.31	0.62	2.83	0.69	18 765.91	4322.39	4.21	0.33
	Impersonal dilemmas ($N = 17$)	4.20	0.77	3.84	0.89	15 714.29	3200.06	4.24	0.41
English/foreign language	Personal dilemmas ($N = 22$)	6.33	0.73	2.68	0.77	34 660.91	5978.32	4.89	0.33
	Impersonal dilemmas ($N = 17$)	4.92	1.08	3.84	1.00	28 819.34	4628.91	4.59	0.31

†The means and standard deviations were computed by the trials that were fully understood by participants. ‡Rating scales of arousal and valence were ranged from 1 to 9. §Rating scale of imagery vividness was ranged from 1 to 7.

prior to making moral choices. We computed their mean arousal and valence ratings, as well as the mean reading time and imagery vividness, of each dilemma ($N = 72$ for both Chinese and English groups, see Table 2).² In all analyses, responses from the participants who did not fully understand the dilemma (i.e. those trials that they failed in the comprehension checks) were excluded. Across all dilemmas, the medians of the percentage of the trials being excluded in the analyses were 16.67% (range: 5.6%–55.6%) and 16.67% (range: 4.2%–69.4%) in Chinese and English groups, respectively.

The present data have a multilevel structure with multiple moral choice (Level 1 within-person) nested within participants (Level 2 between-person). In total, there were 4434 observations collected from 144 participants (i.e. approximately 31 moral choices on average were made by each participant). To account for the within-cluster dependency, a design-based approach was adopted to handle the multilevel data by adjusting for parameter estimate standard errors through the sandwich estimator (Asparouhov, 2005; Satorra & Muthén, 1995). The robust maximum likelihood parameter estimator was used to maximize the loglikelihood function under non-normality data, non-independent observations and binary outcome variable. The four impersonal dilemmas (Donation, Environmental Policy A1, Environmental Policy A2 and Vaccine Policy) did not have a clear utilitarian vs deontological distinction regarding their resolutions. Nonetheless, whether or not they were included in the analyses did not significantly alter the results, so we report the analyses that included these dilemmas.

To test Hypothesis 1, the binary outcome of moral choice (where 1 indicates utilitarian choice, 51.1%, and 0 indicates deontological choice, 48.9%) was regressed on three predictors, namely (i) dilemma type (where 1 indicates personal dilemma and 0 indicates impersonal dilemma), (ii) language (where 1 indicates dilemma in Chinese and 0 indicates dilemma in English) and (iii) arousal. Four covariates, namely valence, dilemma reading time, participants' previous knowledge about the dilemmas and the imagery vivid-

ness of the dilemmas were controlled. First, dilemma type significantly predicted the moral choices made by the participants, B (nonstandardized regression coefficient) = -0.37 , β (standardized regression coefficient) = -0.10 , $p < 0.001$, OR (odds ratio) = 0.69, OR 95% CIs (confidence interval) [0.60, 0.79], indicating that personal dilemmas yielded more deontological choices whereas impersonal dilemmas yielded more utilitarian choices. Second, language had no effect on moral choice, $B = 0.07$, $\beta = 0.02$, $p = 0.39$, OR = 1.08, OR 95% CIs [0.91, 1.27], showing that moral choices were not made systematically in accordance with the use of language. Third, unexpectedly, moral choices were positively predicted by arousal, $B = 0.11$, $\beta = 0.14$, $p < 0.001$, OR = 1.11, OR 95% CIs [1.07, 1.15], revealing that more arousing dilemmas yielded more utilitarian choices, while less arousing dilemmas yielded more deontological choices. In addition, among the covariates, only valence, $B = 0.12$, $\beta = 0.12$, $p < 0.001$, OR = 1.13, OR 95% CIs [1.07, 1.18], and reading time, $B = 0.01$, $\beta = 0.06$, $p = 0.015$, OR = 1.008, OR 95% CIs [1.001, 1.014], positively predicted the moral choices.

To examine Hypothesis 2 that arousal may serve as the mechanism underlying the linkage between dilemma type and moral choice, mediation analysis was conducted. Dilemma type positively predicted arousal, $B = 0.64$, $\beta = 0.13$, $p < 0.001$, 95% CIs [0.47, 0.82], which in turn positively predicted moral choice, $B = 0.11$, $\beta = 0.14$, $p < 0.001$, OR = 1.11, OR 95% CIs [1.07, 1.15]. Given the significant indirect effect, $B = 0.07$, $\beta = 0.02$, $p < 0.001$, OR = 1.07, OR 95% CIs [1.04, 1.10], these results indicate that personal dilemmas may predispose people to make more utilitarian choices than deontological choices through enhancing the perceived arousal of the dilemmas. After controlling for the effect of emotion arousal, the direct effect of dilemma type on moral choice was significant, $B = -0.37$, $\beta = -0.10$, $p < 0.001$, OR = 0.69, OR 95% CIs [0.60, 0.79], and interestingly, the direction of this effect was opposite to the indirect effect, demonstrating a typical case of inconsistent mediation (MacKinnon, Krull & Lockwood, 2000).

To examine Hypotheses 3a and 3b, the binary outcome of moral choice was regressed on language with the control of five covariates, namely arousal, valence, dilemma reading time, participants' previous knowledge about the dilemmas and the imagery vividness of the dilemmas. Among the 22 personal dilemmas (2516 observations nested within 144 participants), results show that moral choices were not made systematically in accordance with the use of language, $B = 0.03$, $\beta = 0.01$, $p = 0.84$, $OR = 1.03$, $OR\ 95\% \text{ CIs } [0.81, 1.30]$. However, to replicate the language effect found by Costa *et al.* (2014), we tested the hypotheses again by confining our analyses to the data of the Footbridge dilemma (111 observations). In that analysis, language negatively predicted moral choice, $B = -1.26$, $\beta = -0.32$, $p = 0.015$, $OR = 0.29$, $OR\ 95\% \text{ CIs } [0.10, 0.79]$, demonstrating that the Chinese group made more deontological choices whereas the English group made more utilitarian choices. Moreover, mediation analysis was conducted to test whether the effect of language on moral choice would work through arousal among the Footbridge dilemma (Hypothesis 3c). After controlling for the effects of valence, dilemma reading time, participants' previous knowledge about the dilemmas and the imagery vividness of the dilemmas, language negatively predicted arousal, $B = -1.06$, $\beta = -0.22$, $p = 0.020$, $95\% \text{ CIs } [-1.96, -0.17]$. However, arousal did not predict moral choice, $B = 0.03$, $\beta = 0.04$, $p = 0.765$, $OR = 1.03$, $OR\ 95\% \text{ CIs } [0.83, 1.29]$, yielding a nonsignificant indirect effect, $B = -0.04$, $\beta = -0.01$, $p = 0.771$, $OR = 0.96$, $OR\ 95\% \text{ CIs } [0.76, 1.22]$.

Auxiliary analyses

It is arguable that the choice made after the emotion ratings may bypass the automatic process and therefore bias the judgement. To address this potential concern, we performed analyses on the two validation groups (one for dilemmas presented in Chinese and one for those presented in English), in which the participants made moral choices immediately after reading the dilemmas without rating the arousal and valence of the dilemmas. The experimental and validation groups were then combined to test whether the requirement of emotion arousal and valence ratings in the experimental groups would distort the subsequent moral judgements. In total, there were 8953 observations nested within 253 participants, moderation analyses were conducted to examine whether the effects of dilemma type, language, dilemma reading time, participants' previous knowledge about the dilemmas and imagery vividness of the dilemmas differed between the experimental and validation groups. (As the participants in the validation groups did not provide emotion arousal and valence ratings, these two variables could not be included in this multilevel analysis.) Overall, all main effects or interactions associated with the 'group status' (experimental vs validation) were nonsig-

nificant ($ps > 0.05$). Thus, the findings we reported above were not significantly compromised by the requirement to perform emotion ratings prior to making moral choices.

Discussion

Previous studies have proposed and supported the dual-process theory of moral judgement (i.e. the dilemma-type effect on moral choice, e.g. Greene *et al.*, 2001, 2004, 2008). Some found a language effect on moral judgement in personal dilemmas (e.g. Costa *et al.*, 2014). Among them, there was a common ground that these effects were related to emotion. Our study measured the self-rated emotion triggered by the dilemmas and tested whether emotion arousal contributed to moral choice and mediated the relationship between dilemma type and moral choice, and between language and moral choice.

Multilevel analyses were conducted to test the hypotheses. The prediction that personal dilemmas yielded fewer utilitarian choices and impersonal dilemmas yielded more utilitarian choices was confirmed, although the mediation analysis showed that the indirect effect of emotion arousal was opposite to the direct effect unexpectedly. We expected that dilemma type would have a positive effect on emotion arousal such that personal dilemmas were rated as more arousing, and emotion arousal would have a negative effect on moral choice, such that higher arousal lead to fewer utilitarian choices. However, we found that the relationship between emotion arousal and moral choice was positive such that higher emotion arousal lead to *more* utilitarian choices. The mediating role of arousal suggests that some participants chose utilitarian choices in personal dilemmas because they felt more aroused. According to Moore *et al.* (2011), making utilitarian choices in personal dilemmas involves more conflict than making deontological choices. It is possible to attribute the more utilitarian choices being made to personal dilemmas to higher arousal. This indirect effect (through arousal) is in contrast to the negative direct effect that personal dilemmas yielded more deontological choices in the current model. Despite the complication, our results suggest that arousal alone may not explain why participants made more deontological choices in personal dilemmas. Future studies should test the dual-process theory further by identifying other emotion factors that might drive the effect of dilemma type on moral choice.

With respect to the foreign language effect, we did not find any significant relationship between language and moral choice, whether a larger set of 39 moral dilemmas or only 22 personal dilemmas were included in the analyses. This suggests that language did not significantly influence moral choice, which could not be explained by the emotion-reducing or cultural-priming effects of language. However,

when we only analyzed the data regarding the Footbridge dilemma that was tested in Costa *et al.*'s (2014) study, we found that participants made more deontological choices than utilitarian choices when they were presented in the native language (Chinese) but not in the foreign language (English), replicating Costa *et al.*'s findings. (This result could not be accommodated by the cultural-priming effect of language.) A recent study (Geipel *et al.*, 2015) also did not find any foreign language effect in dilemmas other than Footbridge dilemma. The absence of language effect in our full dataset suggests that the results obtained in previous research (e.g. Costa *et al.*, 2014) could not be generalized to different moral dilemmas.

As there was a language effect on the judgement in the Footbridge dilemma, the mediating role of emotion arousal was examined. Our findings show that the association between language and moral choice was not mediated by arousal, which is inconsistent with the hypothesis of the emotion-reducing effect of foreign language suggested by Costa *et al.* (2014). Nevertheless, it is important to note that in the present study dilemmas presented in English yielded higher arousal ratings than those presented in Chinese (5.72 and 4.82, respectively, $F(1,76) = 15.15$, $p < 0.001$), which contradicts Costa *et al.*'s premise that dilemmas are more aroused in a native language than in a foreign language. One possibility for this is that the dilemmas were originally written in English and may have lost their imagery vividness when they were translated into Chinese. Indeed, we found that English dilemmas were rated as more vivid than Chinese dilemmas (4.76 and 4.22, respectively, $F(1,76) = 44.27$, $p < 0.001$). Another possibility is that, because students in Hong Kong learn and speak English in classroom contexts and they often link English with their language competence, such that when participants read the moral dilemmas in English, they may experience foreign language anxiety (Horwitz, Horwitz & Cope, 1986), and in turn, rate a generally higher arousal when they read the moral dilemmas in English than in Chinese.

The finding that English dilemmas were rated as more arousing than Chinese dilemmas shows that bilinguals might not necessarily rate dilemmas in a native language to be more arousing than those in a foreign language. It is noteworthy that the nonsignificant indirect effect of the mediation analysis indicates that the language effect on moral choice was not due to arousal ratings, so the unexpected arousal difference between English and Chinese dilemmas *per se* might not account for the absence of language effect on moral choices when the full set of dilemmas is considered. Nevertheless, it is possible that the arousal ratings in the current study reflect the arousal induced by the moral dilemmas, as well as by the language in which the dilemmas were presented. Future studies should tease apart the effect of these two types of emotion arousal and further examine the role of emotion arousal in

the relationship between language and moral choice in the moral dilemma task.

Since we did not find any language effect on moral choice when examining a larger set of moral dilemmas, we cannot draw a specific conclusion on whether our manipulation of foreign vs native language could prime different culture values in Chinese–English bilinguals. It is possible that the emotion-reducing effect of foreign language might overshadow the cultural-priming effect of language for bilingual/bicultural participants. Future research could address this issue by improving the salience of cultural manipulation, so as to place in a more competing position the two predictions derived from the emotion-reducing effect of foreign language and the cultural priming effect of language. For example, Hong Kong students relative to North American students were found to focus more on collective duties than individual rights when they were primed to think about their collective self by completing Chinese sentences starting with 'We being Chinese' (Hong & Chiu, 2001; Hong *et al.*, 2001). Future studies could simultaneously manipulate the language bilingual participants use to make moral judgements and the currently salient identity evoked in bilingual participants.

Conclusion

In the current study, we examined the effect of dilemma type, language and emotion arousal of moral dilemmas and tested the mediating role of arousal on the association between dilemma type and moral choice (i.e. dual-process theory), and between language and moral choice. While we did obtain a direct effect of dilemma type on moral choices in which participants made more deontological choices in personal dilemmas, we also found an inconsistent mediating role of arousal for the indirect effect of dilemma type on moral choices: participants made utilitarian choices in personal dilemmas due to the higher arousal. Regarding the language effect, language failed to predict moral choice when a larger set of moral dilemmas was analyzed, which could not be explained by the emotion-reducing or cultural-priming effects of language. When we focused on the Footbridge dilemma, however, we found that Chinese–English bilinguals made fewer utilitarian choices when it was presented in the native language (Chinese) than when it was presented in the foreign language (English). However, this relationship was not mediated by arousal, suggesting that the significant language effect in the Footbridge dilemma was not due to the emotion-reducing effect of foreign language. Future studies should further identify and test other emotion factors that potentially mediate the relationship between dilemma type and moral choices and the relationship between language and moral choices.

Acknowledgements and funding

There is no financial support or relationships that may pose a conflict of interest associated with the current research.

End notes

1. In this article, 'choice' refers to a participant's decision outcome, whereas 'judgement' refers to the general process in

making a choice (Einhorn & Hogarth, 1981). We used 'moral choice' instead of 'moral judgement' as our participants responded to the dilemmas by choosing one of the two resolutions (a utilitarian and a deontological action).

2. The detailed normative value of experimental and validation groups (e.g. arousal rating, valence rating, and percentages of utilitarian and deontological choices) are available in the supporting information.

References

- Asparouhov, T. (2005). Sampling weights in latent variable modeling. *Structural Equation Modeling*, 12, 411–434. doi:10.1207/s15328007sem1203_4.
- Bradley, M. M. & Lang, P. J. (1994). Measuring emotion: The Self-Assessment Manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25, 49–59. doi:10.1016/0005-7916(94)90063-9.
- Briley, D. A., Morris, M. W. & Simonson, I. (2005). Cultural chameleons: Biculturals, conformity motives, and decision making. *Journal of Consumer Psychology*, 15, 351–362. doi:10.1207/s15327663jcp1504_9.
- Caldwell-Harris, C. L. (2015). Emotionality differences between a native and foreign language implications for everyday life. *Current Directions in Psychological Science*, 24, 214–219. doi:10.1177/0963721414566268.
- Carmona-Perera, M., Martí-García, C., Pérez-García, M. & Verdejo-García, A. (2013). Valence of emotions and moral decision-making: Increased pleasantness to pleasant images and decreased unpleasantness to unpleasant images are associated with utilitarian choices in healthy adults. *Frontiers in Human Neuroscience*, 7, 1–7. doi:10.3389/fnhum.2013.00626.
- Chiu, C. Y., Dweck, C. S., Tong, J. Y. Y. & Fu, J. H. Y. (1997). Implicit theories and conceptions of morality. *Journal of Personality and Social Psychology*, 73, 923–940. doi:10.1037/0022-3514.73.5.923.
- Chiu, C. Y., Leung, A. K. & Kwan, L. (2007). Language, cognition, and culture: Beyond the Whorfian hypothesis. In S. Kitayama & D. Cohen (Eds.), *Handbook of cultural psychology* (pp. 668–688). New York, NY, US: Guilford Press.
- Christensen, J. F., Flexas, A., Calabrese, M., Gut, N. K. & Gomila, A. (2014). Moral judgment reloaded: A moral dilemma validation study. *Frontiers in Psychology*, 5, 1–18. doi:10.3389/fpsyg.2014.00607.
- Cipolletti, H., McFarlane, S. & Weissglass, C. (2015). The moral foreign-language effect. *Philosophical Psychology*, Advance online publication. doi:10.1080/09515089.2014.993063.
- Costa, A., Foucart, A., Hayakawa, S., Aparici, M., Apesteguia, J., Heafner, J., et al. (2014). Your morals depend on language. *PLoS ONE*, 9, e94842. doi:10.1371/journal.pone.0094842.
- Earle, R. A. (1969). Use of the structured overview in mathematics classes. In H. L. Herber & P. L. S. Sanders (Eds.), *Research in reading in the content areas: First year report* (pp. 49–58). NY: Syracuse University Press.
- Einhorn, H. & Hogarth, R. (1981). Behavioral decision theory: Processes of judgement and choice. *Annual Review of Psychology*, 32(11), 53–88. doi:10.1146/annurev.ps.32.020181.000413.
- Foot, P. (1967). The problem of abortion and the doctrine of the double effect. In P. Foot (Ed.), *Virtues and vices and other essays in moral philosophy* (pp. 19–31). Oxford: Blackwell.
- Geipel, J., Hadjichristidis, C. & Surian, L. (2015). How foreign language shapes moral judgment. *Journal of Experimental Social Psychology*, 59, 8–17. doi:10.1016/j.jesp.2015.02.001.
- Greene, J. D. (2009). The cognitive neuroscience of moral judgment. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (4th ed., pp. 987–999). Cambridge, MA: MIT Press.
- Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E. & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, 107, 1144–1154. doi:10.1016/j.cognition.2007.11.004.
- 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. doi:10.1016/j.neuron.2004.09.027.
- 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. doi:10.1126/science.1062872.
- Hong, Y. Y. & Chiu, C. Y. (2001). Toward a paradigm shift: From cross-cultural differences in social cognition to social-cognitive mediation of cultural differences. *Social Cognition*, 19, 181–196. doi:10.1521/soco.19.3.181.21471.
- Hong, Y. Y., Ip, G., Chiu, C. Y., Morris, M. W. & Menon, T. (2001). Cultural identity and dynamic construction of the self: Collective duties and individual rights in Chinese and American cultures. *Social Cognition*, 19, 251–268.
- Hong, Y. Y., Morris, M. W., Chiu, C. Y. & Benet-Martinez, V. (2000). Multicultural minds: A dynamic constructivist approach to culture and cognition. *American Psychologist*, 55, 709. doi:10.1037/0003-066X.55.7.709.
- Horwitz, E. K., Horwitz, M. B. & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70, 125–132. doi:10.2307/327317.
- Keysar, B., Hayakawa, S. L. & An, S. G. (2012). The foreign-language effect thinking in a foreign tongue reduces decision biases. *Psychological Science*, 23, 661–668. doi:10.1177/0956797611432178.
- Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., et al. (2007). Damage to the prefrontal cortex increases utilitarian moral judgments. *Nature*, 446, 908–911. doi:10.1038/nature05631.
- Lotto, L., Manfrinati, A. & Sarlo, M. (2014). A new set of moral dilemmas: Norms for moral acceptability, decision times, and emotional salience. *Journal of Behavioral Decision Making*, 27, 57–65. doi:10.1002/bdm.1782.

- MacKinnon, D. P., Krull, J. L. & Lockwood, C. M. (2000). Equivalence of the mediation, confounding, and suppression effect. *Prevention Science*, 1, 173–181. doi:10.1023/A:1026595011371.
- Mann, S. K. & Cheng, V. (2013). Responding to moral dilemmas: The roles of empathy and collectivist values among the Chinese. *Psychological Reports*, 113, 107–117. doi:10.2466/17.21.PR0.113x14z6.
- Markus, H. R. & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98(2), 224–253. doi:10.1037/0033-295X.98.2.224.
- Moore, A. B., Clark, B. A. & Kane, M. J. (2008). Who shalt not kill? Individual differences in working memory capacity, executive control, and moral judgment. *Psychological Science*, 19, 549–557. doi:10.1111/j.1467-9280.2008.02122.x.
- Moore, A. B., Lee, N. L., Clark, B. A. & Conway, A. R. (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.
- Moretto, G., Làdavas, E., Mattioli, F. & Di Pellegrino, G. (2010). A psychophysiological investigation of moral judgment after ventromedial prefrontal damage. *Journal of Cognitive Neuroscience*, 22, 1888–1899. doi:10.1162/jocn.2009.21367.
- Oyserman, D. & Lee, S. W. S. (2007). Priming ‘culture’: Culture as situated cognition. In S. Kitayama & D. Cohen (Eds.), *Handbook of cultural psychology* (pp. 255–279). New York, NY, US: Guilford Press.
- Oyserman, D. & Lee, S. W. S. (2008). Does culture influence what and how we think? Effects of priming individualism and collectivism. *Psychological Bulletin*, 134, 311–342. doi:10.1037/0033-2909.134.2.311.
- Pastötter, B., Gleixner, S., Neuhauser, T. & Bäuml, K.-H. T. (2013). To push or not to push? Affective influences on moral judgment depend on decision frame. *Cognition*, 126, 373–377. doi:10.1016/j.cognition.2012.11.003.
- Pizarro, D. (2000). Nothing more than feelings? The role of emotions in moral judgment. *Journal for the Theory of Social Behaviour*, 30, 355–375. doi:10.1111/1468-5914.00135.
- Sachdeva, S., Singh, P. & Medin, D. (2011). Culture and the quest for universal principles in moral reasoning. *International Journal of Psychology*, 46, 161–176. doi:10.1080/00207594.2011.568486.
- Sapir, E. (1949). *Selected writings of Edward Sapir*. Los Angeles: University of California Press.
- Satorra, A. & Muthén, B. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, 25, 267–316. doi:10.2307/271070.
- Schneider, W., Eschman, A. & Zuccolotto, A. (2002). *E-Prime reference guide*. Pittsburgh, PA: Psychology Software Tools.
- Thomson, J. J. (1985). The trolley problem. *Yale Law Journal*, 94, 1395–1415.
- Triandis, H. C., McCusker, C. & Hui, C. H. (1990). Multimethod probes of individualism and collectivism. *Journal of Personality and Social Psychology*, 59, 1006–1020. doi:10.1037/0022-3514.59.5.1006.
- Ugazio, G., Lamm, C. & Singer, T. (2012). The role of emotions for moral judgments depends on the type of emotion and moral scenario. *Emotion (Washington, D.C.)*, 12, 579. doi:10.1037/a0024611.
- Whorf, B. L. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. New York: Wiley.
- Young, L., Bechara, A., Tranel, D., Damasio, H., Hauser, M. & Damasio, A. (2010). Damage to ventromedial prefrontal cortex impairs judgment of harmful intent. *Neuron*, 65, 845–851. doi:10.1016/j.neuron.2010.03.003.
- Young, L. & Koenigs, M. (2007). Investigating emotion in moral cognition: A review of evidence from functional neuroimaging and neuropsychology. *British Medical Bulletin*, 84, 69–79. doi:10.1093/bmb/ldm031.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher’s website:

File S1 Normative values.

File S2 List of moral dilemmas.