Accepted Manuscript

Please cite this article as: Geipel, J., Hadjichristidis, C. & Surian, L., How Foreign Language Shapes Moral Judgment, *Journal of Experimental Social Psychology* (2015), doi: 10.1016/j.jesp.2015.02.001

How Foreign Language Shapes Moral Judgment

Janet Geipel, Constantinos Hadjichristidis, Luca Surian

University of Trento

Author Note

Janet Geipel and Luca Surian, Department of Psychology and Cognitive Sciences, University of Trento; Constantinos Hadjichristidis, Department of Economics and Management, University of Trento, and Research Centre for Decision Making, Leeds Business School, Leeds University.

Correspondence concerning this article should be addressed to Janet Geipel, Department of Psychology and Cognitive Sciences, University of Trento, Corso Bettini 31, 38068 Rovereto, Italy or to Constantinos Hadjichristidis, Department of Economics and Management, University of Trento, Via Inama 5, 38122 Trento, Italy. E-mails: janet.geipel@unitn.it; k.hadjichristidis@unitn.it

Abstract

We investigated whether and how processing information in a foreign language as opposed to the native language affects moral judgments. Participants judged the moral wrongness of several private actions, such as consensual incest, that were depicted as harmless and presented in either the native or a foreign language. The use of a foreign language promoted less severe moral judgments and less confidence in them. Harmful and harmless social norm violations, such as saying a white lie to get a reduced fare, were also judged more leniently. The results do not support explanations based on facilitated deliberation, misunderstanding, or the adoption of a universalistic stance. We propose that the influence of foreign language is best explained by a reduced activation of social and moral norms when making moral judgments.

Keywords: moral judgment, foreign language, bilingualism, emotion, social and cultural norms

How Foreign Language Shapes Moral Judgment

Imagine reading about the case of a brother and sister who have an incestuous relationship. What would your moral reaction be? Most people judge incest as wrong, even in circumstances where potential harm is minimized (Haidt, 2001). Now imagine reading the same story in a foreign language that you comprehend well. Would your moral reaction change? It shouldn't – the story is the same (principle of description invariance [Tversky & Kahneman, 1981] or extensionality [Arrow, 1982]).[1] But psychological research on moral violations suggests that it might: A higher proportion of participants judge that it is acceptable to shove a man into the path of a trolley to save five lives, when the scenario and questions are printed in a foreign language rather than in their native language (Cipolletti, McFarlane & Weissglass, 2015; Costa et al., 2014; Geipel, Hadjichristidis, & Surian, 2014).

Here we aimed to extend the foreign language effect to actions that are relatively harmless, but nevertheless typically condemned. We expected that foreign language would distance participants from intuition and gut-feelings, and through that promote less harsh moral judgments. We considered two competing hypotheses. Costa et al. (2014) proposed that a "muted" intuition could make the moral machinery switch from the default automatic, intuitive mode, to a controlled mode, thus focusing the attention to the harmless consequences (see also Cipolletti et al., 2015). We call this the *controlled-processing* hypothesis. Alternatively, the moral machinery might remain on the automatic, intuitive mode, but the muted intuition would nevertheless promote less harsh moral judgments. We call this the *automatic-processing* hypothesis. This could happen either through an attenuation of the typical aversive reaction (see the *affect heuristic*; Slovic, Finucane, Peters, & MacGregor, 2002; Kahneman & Frederick, 2002), or a reduction of the mental accessibility of moral and social rules (e.g., Bond & Lai, 1986; Dewaele, 2010). Moral and

social rules are learned and experienced through interactions involving the native language, and so a foreign language might evoke them to a lesser extent (see Marian & Neisser, 2000).

Prior research

Foreign language has been shown to attenuate emotional response to words and phrases (for reviews see Caldwell-Harris, 2014; Pavlenko, 2012). For example, Harris and colleagues found that childhood reprimands, such as "Don't do that!", evoked reduced skin conductance responses when they were read aloud in a foreign language (Harris, Ayçiçegi, & Gleason, 2003; Harris, Gleason, & Ayçiçeği, 2006). Moreover, a large questionnaire-based study has shown that late bilinguals rated taboo words and swearwords as less emotional in a foreign language than in a native language (Dewaele, 2004; Pavlenko, 2004). Studies also suggest that a foreign language facilitates people to discuss topics that are considered off-limits or taboo in their native language. For example, Bond and Lai (1986) found that Chinese-English bilinguals spoke longer about embarrassing topics, such as sexual attitudes, in a foreign language. In the same vein, Dewaele (2010) found that several UK-based multilinguals preferred using swearwords in a foreign language, stating that a foreign language allows them to escape from social and cultural restrictions. However, some studies have failed to detect an attenuation of emotions (e.g., Ayçiçegi-Dinn, & Caldwell-Harris, 2009; Eilola, Havelka, & Sharma, 2007; Sutton, Altarriba, Gianico, & Basnight-Brown, 2007). To reconcile these findings, Harris and colleagues proposed that the relative emotionality of a foreign versus a native language depends on a complex interplay between age of acquisition, level of proficiency, and the emotional context in which the foreign language is learned and used (Caldwell-Harris, 2014; Harris et al., 2006).

Foreign language has also been shown to reduce decision biases that are believed to have an emotional basis (Keysar, Hayakawa, & An, 2012). Moreover, recent studies demonstrated that it also influences moral judgment (Costa et al., 2014; Geipel et al. 2014).

This research was confined to trolley dilemmas (Foot 1978; Thompson, 1985) that create tension between a characteristically utilitarian perspective, which aims at maximizing net benefit, and a characteristically deontological perspective, which forbids actions that harm innocent others. You are informed that a runaway trolley will kill five people unless an action is performed, either pulling a lever (standard trolley dilemma) that would make the trolley switch to alternative tracks where one workman is standing, or by pushing a person off a bridge (footbridge dilemma). Is it morally acceptable to perform such actions? Adults and children by the age of four typically respond that it is acceptable to pull the lever, but not to push the person (Cushman, Young, & Hauser, 2006; Pellizzoni, Siegal, & Surian, 2010).

The dual-process theory of moral judgment (e.g., Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001) explains these findings by suggesting that responses result from a competition between an automatic, emotional system that prompts a deontological response, and a slow, controlled system that favors a utilitarian response.[2] When the proposed action is emotionally salient (pushing a person off a bridge), the emotional system predominates; when it is not (pulling a lever), the controlled system overrides the emotional system and produces a characteristically utilitarian response (see also Koenigs et al., 2007).

When these trolley dilemmas were presented in a foreign language, utilitarian responses increased but just for the footbridge dilemma (Costa et al., 2014; Geipel et al., 2014). Furthermore, as proficiency in the foreign language increased, language differences decreased. These results were robust across a variety of foreign—native language combinations and cultures (for a replication, see Cipolletti et al., 2015). The proposed explanation is that foreign language triggers emotional distance, which privileges controlled processing (controlled-processing hypothesis). Its effects are observed in the footbridge

dilemma, as this is typically processed by the emotional system, but not in the trolley dilemma, which is commonly processed by the controlled system (Greene et al., 2001).

Notice that these findings are also compatible with the automatic-processing hypothesis. The footbridge dilemma involves a prohibited action (pushing a person; see Cushman, 2013), whereas the trolley dilemma does not. It could be that foreign language promoted utilitarian responses for the footbridge dilemma, because it allowed people to see past the taboo action (either by reducing the aversive response linked to the prohibition, and/or by deactivating social and moral norms). This interpretation is consistent with recent research that shows that characteristically utilitarian responses do not necessarily imply controlled processes, but may also arise from impaired social cognition, such as reduced empathy (see Duke & Bègue, 2015; Kahane, Everett, Earp, Farias, & Savulescu, 2015).

Previous studies examining the role of foreign language on moral judgment have four limitations. First, they have examined only the trolley dilemmas, which involve severe personal harm and concern contrived cases distant from the participants' experience (Hare, 1981; Sunstein, 2005). Second, these dilemmas involve a numerical tradeoff (killing *one* in order to save *five*). As processing information in a foreign language is difficult, people might have treated the dilemmas as simple math problems (Bloom, 2011). Third, these studies offer no empirical support for the central claim that language has a cooling effect on emotions, or that this cooling effect prompts controlled (utilitarian) reasoning. Fourth, the results are open to an in-group out-group interpretation (Caldwell-Harris, 2014). Participants reading the materials in a foreign language might have inferred that the scenarios concerned foreign people (out-group), whereas those reading them in the native language might have inferred that they concerned co-nationals (in-group). Research suggests that feeling socially connected to the characters portrayed in a scenario influences moral judgment (e.g., Bloom, 2011; Greene, 2013; Lucas & Livingston, 2014). Thus, the observed foreign language effect might

reflect added assumptions, rather than the use of foreign language per se. In the present study we address all these issues.

Present research

Our first aim was to broaden the scope of the foreign language effect on moral judgment. We examined different types of violations that, according to the categorization proposed by Shweder, Much, Mahapatra and Park (1997; see also Guerra & Giner-Sorolla, 2010), concern the ethics of Community (e.g., violations of loyalty), Autonomy (e.g., violations of fairness) and Divinity (e.g., violations of purity) (CAD for short; for an extension of this model see Graham, Haidt, & Nosek, 2009; Haidt & Joseph, 2008). We selected violations that did not involve physical harm, such as siblings having consensual and safe sex (see e.g., Björklund, Haidt, & Murphy, 2000; Eyal, Liberman, & Trope, 2008; adapted from Haidt, 2001). People typically judge such behaviors as ethically wrong, but struggle to supply moral justifications (*moral dumbfounding*; Haidt, Koller, & Dias, 1993). To test the generalizability of this effect, we also asked participants to evaluate relatively harmful and harmless social norm violations in community and autonomy ethics. We predicted that foreign language would promote less harsh moral judgments.

Our second aim was to test whether the effect of foreign language on moral judgment is underpinned by an attenuation of emotions, as previous studies have suggested. A third aim was to differentiate between the *automatic-* and *controlled-processing* hypotheses. To this purpose, we used two tasks. We asked participants to state their confidence in their moral evaluations, and to answer a tricky question (see Study 3), which tests the ability to override an intuitive wrong answer. Finally, we assessed a number of deflationary explanations of the foreign language effect, such as that it is simply due to misunderstanding, or in-group out-group considerations.

Study 1

In Study 1, we examined whether reading moral transgressions in a foreign versus a native language influences moral wrongness judgments (see Table 1 for descriptions of the scenarios). We tested native German speakers and native Italian speakers. For both, the foreign language was English. We predicted that the use of a foreign language would promote less harsh moral judgments.

Methods

Participants.

Study 1a. Forty-eight students (34 female, 14 male; mean age = 27.27 years, range: 18–70 years) from the Free University of Berlin participated at the beginning of a lecture.[3] Twenty-seven students were randomly assigned to the foreign language condition and received a questionnaire in English and 21 to the native language condition and received its German equivalent. Ten participants were excluded from the analyses, as they were not native German speakers. For all studies, details of the participants assigned to the foreign language conditions are presented in Appendix A.

Study 1b. Sixty-four students (56 female, 6 male, 2 unknown; mean age = 20.56 years, range: 19–24 years) from the University of Trento participated at the beginning of an English class for credits.[4] Thirty-six were randomly assigned to the foreign language condition (English), and 28 to the native language condition (Italian). Four participants were excluded from the analyses, as they were not native Italian speakers.

Materials and Procedure. We used four scenarios (adapted from Haidt et al., 1983; Eyal et al., 2008) each describing one or more persons committing a moral violation (see Table 1, dog, incest, exam and flag items). Participants were asked to judge the wrongness of each action on a scale ranging from 0 (*perfectly ok*) to 9 (*extremely wrong*). Two presentation orders were created and counterbalanced across participants. In all our studies the original materials were in English, they were translated by individuals who are highly proficient in the

foreign and native languages used, and checked by bilingual speakers for comparability. The two versions of each scenario were also closely matched for word counts.

- -
Table 1 about here
- -

Results

Study 1a. The results are illustrated in the left panel of Figure 1. We submitted the mean wrongness ratings to a 2 (Language: foreign vs. native) × 4 (Scenario: 1-4) mixed factor analysis of variance (ANOVA), with repeated measures on scenario. As predicted, there was a significant main effect of language, $F(1, 34) = 6.44$, $p = .016$, $f = .43$. The scenarios were judged less harshly in the foreign language ($M = 3.57$, CI [2.83, 4.31])[5] than in the native language ($M = 4.91$, CI [4.13, 5.70]). There was also a significant main effect of scenario, $F(3, 102) = 24.52$, $p < .001$, $f = .85$. Mean wrongness ratings for the dog scenario ($M = 6.14$, CI [5.21, 7.07]) and the incest scenario ($M = 5.53$, CI [4.57, 6.49]) were significantly higher than those for the exam scenario ($M = 3.42$, CI [2.40, 4.44]) and the flag scenario ($M = 1.72$, CI [0.95, 2.50]). There was no language × scenario interaction, $F(3, 102) = 1.34$, $p = .265$, $f = .20$.

Study 1b. The results are illustrated in the right panel of Figure 1. We submitted the mean wrongness ratings to a 2 (Language: foreign vs. native) × 4 (Scenario: 1-4) mixed-factor ANOVA, with repeated measures on scenario. As in Study 1a, we found a significant main effect of language, $F(1, 58) = 4.93$, $p = .030$, $f = .29$. The scenarios were judged less harshly when presented in the foreign language ($M = 5.40$, CI [4.73, 6.07]) than in the native language ($M = 6.55$, CI [5.76, 7.34]). There was a significant main effect of scenario, $F(3, 174) = 15.57$, $p < .001$, $f = .52$. The mean wrongness ratings of the incest scenario ($M = 6.93$, CI [6.23, 7.64]), the dog scenario ($M = 6.20$, CI [5.38, 7.02]), and the exam scenario ($M = 6.40$, CI [5.77, 7.03]) were all significantly higher than the mean wrongness rating of the flag

scenario ($M = 4.36$, CI [3.67, 5.05]). No other differences were observed. There was no

language × scenario interaction, $F(3, 174) = 0.54$, $p = .656$, $f = .10$.

- -

Figure 1 about here

- -

In sum, Studies 1a and 1b show that the foreign language effect on moral judgment

generalizes to private violations involving relatively harmless but offensive consequences. As

anticipated, the use of a foreign language promoted less harsh moral judgments.

Study 2

In Study 2, we presented a new sample of late Italian-English bilinguals with the same

four scenarios. In addition to moral judgments, we also asked participants to rate their

emotional reactions. Here, our main aim was to examine whether foreign language influences

moral judgments by attenuating emotions.

Methods

Participants. Seventy-eight English majors (61 female, 15 male, 2 unknown; mean

age = 23.11 years, age range: 20–38 years) from the University of Verona participated at the

beginning of an English lesson; 42 were randomly assigned to the foreign language condition

(English), and 36 to the native language condition (Italian).[6]

Materials and Procedure. Following each scenario, participants judged the

wrongness of the action that was depicted in it. They were also asked to rate how *upset*,

worried, *disgusted*, *sad*, and *angry* they felt while reading the scenario using a 5-point scale

(1 = *very slightly or not at all*, 2 = *a little*, 3 = *moderately*, 4 = *quite a bit*, 5 = *extremely*; from

Watson, Clark, & Tellegen, 1988). The presentation order of the moral judgment and emotion

rating tasks was counterbalanced. Preliminary analyses revealed no order effects and so we

dropped this factor from the analyses.

Results and discussion

Emotion ratings. The five emotion scales were highly associated (Cronbach's alpha = .85 in the native language condition, Cronbach's alpha = .85 in the foreign language condition). Thus, we reduced the five scales into an emotion index score by taking their average (see Figure 2). The resulting scores were submitted to a 2 (Language: foreign vs. native) × 4 (Scenario: 1-4) mixed-factor ANOVA, with repeated measures on scenario. Although there was no main effect of language, $F(1, 76) = 1.24$, $p = .268$, $f = .13$, there was a significant language × scenario interaction, $F(3, 228) = 5.56$, $p = .001$, $f = .26$. Simple one-way ANOVAs revealed that foreign language attenuated emotions in the dog scenario ($M_{FL} = 3.33$, $M_{NL} = 3.88$), $F(1, 76) = 6.30$, $p = .014$, $f = .29$, and in the incest scenario ($M_{FL} = 2.84$, $M_{NL} = 3.48$), $F(1, 76) = 7.22$, $p = .009$, $f = .31$, but not in the exam scenario ($M_{FL} = 2.26$, $M_{NL} = 2.05$), $F(1, 76) < 1$, $p = .341$, $f = .11$, and flag scenario ($M_{FL} = 2.45$, $M_{NL} = 2.12$), $F(1, 76) = 1.62$, $p = .207$, $f = .15$. There was also a main effect of scenario, $F(3, 228) = 42.89$, $p < .001$, $f = .75$. Post hoc comparisons indicated that the mean emotion ratings of the dog scenario ($M = 3.61$, CI [3.39, 3.83]), the incest scenario ($M = 3.16$, CI [2.92, 3.40]) and the flag scenario ($M = 2.28$, CI [2.03, 2.54]) were significantly higher than the mean emotion rating of the exam scenario ($M = 2.16$, CI [1.93, 2.38]).

- -
Figure 2 about here
- -

Moral judgments. Previous studies have suggested that foreign language influences moral judgment by attenuating emotions. We thus grouped the scenarios into those that showed an attenuation of emotions (dog, incest) and those that did not (exam, flag). We predicted that foreign language would promote less harsh moral judgments but for only the dog and incest scenarios. Furthermore, these items were relatively more emotional than the exam and flag scenarios, and previous research suggests that the effect of foreign language is

confined to high emotion items, such as the footbridge dilemma (Costa et al., 2014; Geipel et

al., 2014).

The findings are illustrated in Figure 3 and are consistent with this prediction. We

analyzed the mean wrongness ratings with simple one-way ANOVAs. For the dog and incest

scenarios foreign language promoted less harsh moral judgments ($M = 7.75$, CI [7.22, 8.27])

than the native language ($M = 8.47$, CI [8.15, 8.79]), $F(1, 76) = 5.17$, $p = .026$, $f = .23$. For the

exam and flag scenarios, foreign language ($M = 5.59$, CI [4.88, 6.15]) and native language ($M

= 5.68$, CI [5.08, 6.33]) induced similar moral judgments, $F(1, 76) = 0.04$, $p = .851$, $f < .01$.

- -
Figure 3 about here
- -

Correlations between emotion and moral judgment ratings. If foreign language

promotes dispassionate controlled thinking, then moral judgments should rely less on

emotions and gut feelings. Therefore the correlation between moral judgments and emotion

ratings should be weaker in the foreign language condition than in the native language

condition (see Shiv & Fedorikhin, 1999). Contrary to this prediction, both correlations were

statistically significant (native language: $r[36] = .70$, $p < .001$; foreign language: $r[41] = .56$,

$p < .001$), and did not differ from one another ($z = 1.01$, $p = .311$).

Mediation analyses. We performed a mediation analysis using the SOBEL macro by

Preacher and Hayes (2004). We used only the dog and incest scenarios, for which we

detected an attenuation of emotions. We used the bootstrapping procedure (5000

bootstrapped re-samples). Figure 4 illustrates the results. There was a significant direct effect

of language on moral judgment ($B = -.72$, $SE = .32$, $p = .026$). The indirect effect controlling

for emotions lies between -1.020 and -0.085 with 99% confidence ($B = -.46$, $SE = .19$).

Because this interval does not include 0, we can conclude that emotion mediates the

association between language and moral judgment.

- -
Figure 4 about here
- -

As a check of the proposed mediation hypothesis, we also conducted an analysis based on a reverse mediation model. We assessed whether the association between language and emotions is mediated by moral judgments. There was a direct effect of language on emotions (B = -.55, SE = .18, p = .003). The indirect effect of language on emotions controlled for moral judgment lies between -0.460 and 0.009 with 99% confidence (B = -19, SE = .09). Since this interval includes 0, we can conclude that there is no reverse mediation. Taken together, the analyses suggest that, for the dog and incest scenarios, emotions mediated the effect of language on moral judgment.

In sum, in Study 2 we found that the use of a foreign language promotes less harsh moral judgments but only for the dog and incest scenarios. For these scenarios, foreign language influenced moral judgment through an attenuation of emotions. Moreover, in both language conditions moral judgments and emotion ratings were significantly correlated and to a similar extent, which is consistent with the automatic-processing hypothesis.

Study 3

In Study 3 we further examined whether foreign language influences moral judgment through an attenuation of emotions. Study 2 found support for this hypothesis but only in two out of four scenarios. Interestingly, these scenarios concerned violations of purity, whereas the other scenarios concerned a violation of fairness (exam) and loyalty (flag). Could it be that the effect of foreign language on moral judgment is mostly confined to purity violations? Studies 1a and 1b, as well as previous studies on trolley dilemmas (Costa et al., 2014; Geipel et al., 2014), suggest that this is not the case, but Study 2 leaves open this possibility. In Study 3 we addressed this question by testing two violations of purity (dog, incest) and two violations of fairness (exam, bonus; see Table 1). As a further test of the generalizability of

the foreign language effect, we also asked participants to rate the moral wrongness of 15

items containing relatively harmful (e.g., *Sell someone a defective car*) and harmless (e.g.,

Fail to vote in minor elections) social norm violations in community and autonomy ethics

(see Appendix B).

   A second aim of Study 3 was to provide evidence to distinguish between the two

competing hypotheses. To this end, we used two new tasks: the Moses illusion task and the

confidence-rating task. In the Moses illusion task (Song & Schwarz, 2008; see Appendix B

for full instructions) participants are asked: "How many animals of each kind did Moses take

on to the ark?" The correct answer is "can't say" (since the biblical character was Noah), but

most people are unable to override the automatic response "two" (Alter, 2013). If foreign

language promotes deliberation, then it should improve performance. The automatic-

processing hypothesis predicts no such improvement. If anything, the increased burden on

cognitive resources might deteriorate performance.

   In the confidence-rating task, participants rated how sure they were in their moral

evaluations. The automatic-processing hypothesis predicts that when using a foreign

language people would be less confident in their judgments because they might lack the "it

feels wrong!" signal that accompanies a strong aversive reaction,[7] which could be grounded

on emotions and/or sociocultural norms. In contrast, the controlled-processing hypothesis

predicts that people would be more confident, as their judgments will be a product of careful

analysis (for further evidence that deliberative thinking leads to higher confidence, see Mata,

Ferreira, & Sherman, 2013).

   In Study 3 we also controlled for the possibility that the foreign language effect is due

to misunderstanding (we asked participants to translate the materials). Furthermore we

assessed whether it is constrained to the use of English, the modern lingua franca, as a

foreign language (here we used German). Participants were told explicitly that the scenarios

took place in their native country and involved co-nationals, in order to rule out possible in-group out-group factors. To assess whether foreign language promotes a universalistic stance, we asked participants to rate how close they feel to *People in my community*, *Italians*, and *People around the world*.[8]

Methods

Participants. Seventy-four German major students (67 females, 7 males; mean age = 21.03 years, age range: 18–30 years) from the University of Trento participated at the beginning of a German lesson.[9] Participants were randomly assigned to either the foreign language condition (German; $n = 37$) or the native language condition (Italian; $n = 37$). Analyses of the translations revealed that two participants assigned in the foreign language condition mistranslated one scenario each. The scores for these two scenarios were excluded from the analyses.

Materials and Procedure. We used two purity violations (dog, incest), two fairness violations (exam, bonus), and two non-moral scenarios (brand, train; see Table 1), both of which should be judged as "perfectly ok." Following a scenario, participants received either the moral judgment task or the emotion-rating task (in counterbalanced order). We used the same scales as in Study 2. Preliminary analyses revealed no effect of order, so we dropped this factor from the analyses. Following the moral judgment task, participants were asked "How sure are you in your evaluation?" and were given a scale ranging from 1 (*not at all sure*) to 7 (*very sure*). Next, participants received the Moses illusion task (see Appendix B), and then a subscale of the *Identification with All Humanity Scale* (McFarland, Webb, & Brown, 2012). Participants were asked: "How close do you feel to each of the following groups?": *People in my community*, *Italians*, *People around the world* and were given a scale ranging from 1 (*not at all close*) to 5 (*very close*). Finally, participants evaluated 15 violations

of everyday moral and social norms on a scale ranging from 1 (*not wrong*) to 4 (*severely wrong*) (see Appendix B).

Results and discussion

Emotion ratings. As in Study 2, the five emotion scales were highly associated (Cronbach's alpha = .92 in the native language condition, Cronbach's alpha = .85 in the foreign language condition). We thus computed an emotion index by taking the mean score over the five scales. These mean emotion scores were submitted to a 2 (Language: foreign vs. native) × 4 (Scenario: 1-4) mixed-factor ANOVA, with repeated measures on scenario. There was no main effect of language, $F(1, 69) = 0.46$, $p = .501$, $f = .08$, but a marginally significant language × scenario interaction, $F(3, 207) = 2.32$, $p = .077$, $f = .18$. Simple one-way ANOVAs revealed that foreign language attenuated emotions for only the dog scenario ($M_{FL} = 3.70$, CI [3.37, 4.03], $M_{NL} = 3.13$, CI [2.73, 3.53], $F(1, 72) = 4.98$, $p = .029$, $f = .26$. As in the previous studies, there was a significant main effect of scenario, $F(3, 207) = 29.79$, $p < .001$, $f = .66$. Post hoc comparisons indicated that the mean emotion rating of the dog scenario ($M = 3.42$, CI [3.16, 3.68]) was significantly higher than the mean emotion ratings of the bonus ($M = 2.91$, CI [2.64, 3.18]) and exam scenarios ($M = 2.06$, CI [1.83, 2.29]). The mean emotion rating of the incest scenario ($M = 3.14$, CI [2.88, 3.39]) was significantly higher than the mean emotion rating of the exam scenario.

We next tested whether proficiency in a foreign language is associated with emotion ratings. We created a proficiency score by aggregating a participant's self-ratings in reading and understanding (both scales ranged from 1 = *almost none*, to 5 = *very good*). The highest possible score is 10, which we also assigned to the participants in the native language condition. We found no significant association between proficiency and mean emotion ratings, $r(72) = .12$, $p = .322$.

Moral judgments. If the influence of foreign language on moral judgments is mediated by an attenuation of emotion, then we should observe no language effect on moral judgment, or perhaps an effect for only the dog scenario. The results of a 2 (Language: foreign vs. native) × 4 (Scenario: 1-4) mixed-factor ANOVA do not support this prediction (see Figure 5). There was a significant main effect of language condition, $F(1, 68) = 8.28$, $p = .005$, $f = .35$, which was not qualified by a language × scenario interaction, $F(3, 204) = 0.67$, $p = .573$, $f = .10$. The scenarios were judged less harshly in the foreign language ($M = 6.77$, CI [6.33, 7.21]) than in the native language ($M = 7.65$, CI [7.23, 8.07]). There was also a significant main effect of scenario, $F(3, 204) = 30.51$, $p < .001$, $f = .66$. The incest scenario received the highest mean moral wrongness rating ($M = 8.15$, CI [7.75, 8.56]), followed by the bonus scenario ($M = 7.92$, CI [7.45, 8.39]), the dog scenario ($M = 7.47$, CI [6.97, 7.97]), and the exam scenario ($M = 5.30$, CI [4.67, 5.93]). Post hoc comparisons indicated that the mean wrongness ratings of the dog scenario, the incest scenario and the bonus scenario were all significantly higher than the mean wrongness rating of the exam scenario.

- -
Figure 5 about here
- -

We also computed a correlation between proficiency in a foreign language and moral judgment ratings. We found that the higher the language proficiency, the harsher the moral judgment: $r(72) = .25$, $p = .034$.

Correlations between emotion and moral judgment ratings. As in Study 2, within each language condition we computed a correlation between moral judgments and (negative) emotion ratings. Both correlations were statistically significant (native language: $r[35] = .49$, $p = .002$; foreign language: $r[35] = .41$, $p = .012$), and not different from one another ($z = -0.39$, $p = .699$). In both languages, higher negative emotion ratings were associated with more severe moral wrongness judgments.

Confidence ratings. The automatic-processing hypothesis predicts that foreign language would decrease confidence in one's moral evaluations, whereas the controlled-processing hypothesis suggests the opposite. The results from a 2 (Language: foreign vs. native) × 4 (Scenario: 1-4) mixed-factor ANOVA support the automatic-processing hypothesis (see Figure 6). There was a significant main effect of language, $F(1, 68) = 9.61$, $p = .003$, $f = .38$, which was not qualified by a language × scenario interaction, $F(3, 204) = 2.35$, $p = .074$, $f = .18$. Participants in the foreign language condition were less confident in their moral judgments ($M = 6.07$, CI [5.81, 6.33]) than participants in the native language condition ($M = 6.63$, CI [6.38, 6.88]). There was also a significant main effect of scenario, $F(3, 204) = 4.41$, $p = .005$, $f = .25$. The incest scenario received the highest mean confidence rating ($M = 6.65$, CI [6.42, 6.88]), followed by the bonus scenario ($M = 6.46$, CI [6.21, 6.71]), the dog scenario ($M = 6.18$, CI [5.90, 6.47]), and the exam scenario ($M = 6.10$, CI [5.78, 6.43]). Post hoc comparisons indicated that the mean confidence rating of the incest scenario was significantly higher than the mean confidence ratings of the dog and exam scenarios.

- -
Figure 6 about here
- -

We then computed a correlation between proficiency in a foreign language and confidence ratings. It was significant and positive: $r(72) = .42$, $p < .001$; the higher the language proficiency, the higher the participants' confidence in their moral judgments.

Moses illusion task. If foreign language promotes analytic reasoning, as the controlled-processing hypothesis claims, then it should increase the frequency of correct responses in this task. It did not. In the native language condition 35.1% of participants responded correctly ("can't say"), compared to 16.2% in the foreign language condition, $\chi^2(1, N = 74) = 3.47$, $p = .062$, $\varphi = -.22$. For the control item, no differences were observed between the two language conditions, $\chi^2(1, N = 74) = 2.38$, $p = .123$, $\varphi = -.24$.

Identification with All Humanity Scale. If foreign language promotes a

universalistic stance, then we should observe differences in terms of how much participants

identify with close and distant social groups. We analyzed the data using a 2 (Language:

foreign vs. native) × 3 (Social group: people in my community vs. Italians vs. people around

the world) ANOVA, with repeated measures on the last factor. There was no main effect of

language: Participants in the foreign language condition gave similar closeness ratings (M =

3.56, CI [3.33, 3.78]) as participants in the native language condition (M = 3.69, CI [3.47,

3.92]), $F(1, 71) = 0.75$, $p = .389$, $f = .10$. Also, there was no main effect of social group, $F(2,$

$142) = 1.05$, $p = .353$, $f = .12$, and no language × social group interaction, $F(2, 142) = 0.90$, p

$= .409$, $f = .11$.

Everyday moral and social norms. Here we addressed whether the foreign language

effect generalizes to the evaluation of norm violations that concern relatively harmless and

harmful actions in community and autonomy ethics. To the extent that it does, foreign

language should promote less harsh moral judgments. We tested this hypothesis by

conducting two analyses of variance, one treating subjects as a random factor (F_1), the other

items (F_2). In accord with the automatic-processing hypothesis, participants in the foreign

language condition gave less harsh moral judgments (M = 2.67, CI [2.54, 2.80]) than

participants in the native language condition (M = 2.94, CI [2.83, 3.05]), $F_1(1, 73) = 10.17$, p

$= .002$, $f = .45$. This result was robust in the analysis by items, $F_2(1, 14) = 11.76$, $p = .004$. In

13 (out of 15) items, the means were in the expected direction (the exact binomial probability

of getting 13 or more hits out of 15 trials is .007, two-tailed). We also computed a correlation

between moral judgments and proficiency. It was significant and positive: $r(72) = .36$, $p =$

.001; the higher the language proficiency, the harsher the moral judgment.

General Discussion

The use of a foreign language, as opposed to a native language, elicited less harsh moral judgments for actions that violate purity, fairness, and loyalty norms, but have relatively harmless consequences. This was true across three native-foreign language combinations: German-English, Italian-English, and Italian-German. The use of a foreign language also elicited less harsh moral judgments for fifteen violations of everyday social and moral norms in community and autonomy ethics. Thus, the present findings consolidate and extend previous ones regarding the trolley dilemmas (Costa et al., 2014; Geipel et al., 2014). Critically, in contrast to previous studies, the present findings are not open to explanations based on misunderstanding, added assumptions concerning who is involved in the scenarios (e.g., in-group or out-group members), a generic bias that distorts the use of the rating scale (the effect was present in scenarios that induce both low and high levels of acceptance, and was absent from non-moral scenarios), or people reducing a moral judgment to a simple math problem.

The present studies provide limited support for the claim that the effect of foreign language is mediated by an attenuation of emotions. Such an effect was found only for two out of four violations in Study 2. In Study 3 we found a main effect of foreign language on moral judgments, but no attenuation of emotions. The failure to detect a widespread attenuation of emotions could be related to how we measured them. Research suggests that emotional scales with verbal anchors (e.g., 1 = *not at all disgusted* to 5 = *extremely disgusted*) elicit higher ratings when the anchors are in a foreign language than in a native language (*the anchor contraction effect*; see de Langhe, Puntoni, Fernandez, van Osselaer, 2011). Presumably emotional anchors are felt less strongly in a foreign language, and thus participants compensate by selecting more extreme ratings. Another potential issue is that some emotion words might lack direct translation equivalents in a foreign language (Pavlenko, 2008). Future research could overcome these issues by eliciting emotions through

emotional scales labeled in the native language, scales that are supplemented by nonverbal cues such as emoticons or colors (see de Langhe et al., 2011, Studies 8 & 9), or by using more direct measures of emotions such as facial affect.

The present findings are not consistent with the idea that foreign language promotes a switch from intuitive to controlled processes (see Keysar et al., 2012; Costa et al., 2014), but rather suggest that intuitive processes remain active (see also Hadjichristidis, Geipel, & Savadori, in press). First, foreign language promoted less confidence in one's moral evaluations. This finding suggests that foreign language makes people judge in accord to weakened or confused intuitions rather than enlightened utilitarian reasoning. An enlightened utilitarian should consider carefully all outcomes and so be confident in his or her decision. Standard economic theory cannot explain certain doubts in a rational agent (see Shafir, Simonson, & Tversky, 1993). Rational agents compute the (expected) utility associated with each option, and choose the one with the highest value (see also Mata et al., 2013). However, an individual who relies on weak intuitions should be much less confident. The reduction in gut feelings that makes this person less radical in his or her moral positions, also explains the reduced confidence in those positions.

Second, foreign language did not improve performance on the Moses illusion task. Costa and colleagues (Costa et al., 2013, Study 4) have similarly failed to find improved performance on the Cognitive Reflection Test (CRT; Frederick, 2005), a logical task where correct responding necessitates the inhibition of an intuitive answer that is incorrect. These authors suggest that the foreign language effect might be confined to problems that have an emotional component. An independent reason to doubt the controlled-processing hypothesis is that thinking in a foreign language should increase cognitive load, and thus hinder rather than facilitate analytical and deliberative reasoning (see Keysar et al., 2012).

One reason why foreign language might promote less harsh moral judgments can be traced to memory and socio-cultural learning processes. The language in which an event is encoded facilitates its recall (e.g., Marian & Neisser, 2000; Schrauf & Rubin, 2000, 2004). All the moral violations we studied concerned norms that have been learned directly or indirectly through social interactions involving the native language. Therefore, a native language is more likely to activate these social and moral norms than a foreign language. In support of this, Gawinkowska and colleagues (Gawinkowska, Paradowski, & Bilewicz, 2013) demonstrated that bilinguals use stronger words to translate swearwords from a native-to-a-foreign language than vice versa, especially for politically incorrect swearwords, such as ones directed at social groups (*ethnophaulisms*). These authors argue that a foreign language exempts bilinguals from self- or socially-imposed norms, thus making them more prone to offending others (see also Bond & Lai, 1986; Dewaele, 2010).

This account helps explaining the present findings: all five moral transgressions as well as the fifteen violations of everyday moral and social norms involved behaviors which have been learned and experienced predominantly in contexts in which the native language was used. It can also explain the findings with the trolley dilemmas. The foreign language effect was present in the footbridge dilemma but absent in the trolley dilemma, because only the footbridge dilemma involves a prohibited action (pushing a person; see also Cushman, 2013). Similarly, the effect was absent from the non-moral dilemmas, because these dilemmas did not involve social norms.

A further possibility is that the use of a foreign language might prompt a generic feeling of uncertainty, which in turn promotes less extreme moral judgments. Here we cannot address this hypothesis as the confidence ratings were tied to the moral evaluations. But future studies could examine whether foreign language also reduces confidence in one's responses concerning emotion-neutral items, such as general knowledge questions.

Conclusion

The present research extends the foreign language effect to harmless-but-offensive actions, but also to relatively harmful and harmless violations of everyday social norms. Foreign language promoted less harsh moral judgments and less confidence in one's moral evaluations. The present findings do not support the view that the use of a foreign language turns people into enlightened rationalists, reasoning coldly in terms of utilitarian principles. Rather, the picture that emerges is of people who are guided by a muted intuition, perhaps due to reduced activation of relevant moral and cultural norms. Whatever the final verdict might be in the theoretical arena, studying how foreign language influences moral judgment is of applied interest, as international public policy involves communicating and processing materials in a foreign language before taking decisions that impact on the populations of many countries.

Reference

Alter, A. L. (2013). The benefits of cognitive disfluency. *Current Directions in Psychological*

  Science, *22*, 437-442. doi:10.1177/0963721413498894

Arrow, K. J. (1982). Risk perception in psychology and economics. *Economic Inquiry*, *20*, 1–

  9.

Ayçiçegi-Dinn, A., & Caldwell-Harris, C. L. (2009). Emotion-memory effects in bilingual

  speakers: A levels-of-processing approach. *Bilingualism: Language and Cognition*, *12*,

  291-303. doi:10.1017/S1366728909990125

Björklund, F., Haidt, J., & Murphy, S. (2000). Moral dumbfounding: When intuition finds no

  reason. *Lund Psychological Reports*, *1*(2), 1-15.

Bloom, P. (2011). Family, community, trolley problems, and the crisis in moral psychology.

  The Yale Review, *99*, 26-43. doi: 10.1111/j.1467-9736.2011.00701.x

Bond, M. H., & Lai, T.-M. (1986). Embarrassment and code-switching into a second

  language. *Journal of Social Psychology*, *126*, 179-186.

Caldwell-Harris, C. L., & Ayçiçeği-Dinn, A. (2009). Emotion and lying in a non-native

  language. *International Journal of Psychophysiology*, *71*, 193-204.

  doi:10.1016/j.ijpsycho.2008.09.006

Caldwell-Harris, C., L., Tong, J., Lung, W., & Poo, S. (2011). Physiological reactivity to

  emotional phrases in Mandarin–English bilinguals. *International Journal of*

  Bilingualism, *15*, 329-352. doi: 10.1177/1367006910379262

Caldwell-Harris, C. (2014, May 20). In a foreign language, "killing 1 to save 5" may be more

  permissible. *Scientific American.* Retrieved from

  http://www.scientificamerican.com/article/in-a-foreign-language-killing-1-to-save-5-

  may-be-more-permissible/

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., & Keysar, B. (2014). Moral judgment depends on language. *PLoS ONE*, *9*, e94842. doi:10.1371/journal.pone.0094842

Costa, A., Foucart, A., Arnon, I., Aparici, M., & Apesteguia, J. (2014). 'Piensa' twice: On the foreign language effect in decision making. *Cognition*, *130*, 236-254. doi:10.1016/j.cognition.2013.11.010

Cushman, F. (2013). Action, outcome and value: A dual-system framework for morality. *Personality and Social Psychology Review*, *17*, 273-292. doi: 10.1177/1088868313495594

Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuitions in moral judgment: Testing three principles of harm. *Psychological Science*, *17*, 1082-1089. doi: 10.1111/j.1467-9280.2006.01834.x

Dewaele, J.-M. (2004). The emotional force of swearwords and taboo words in the speech of multilinguals. *Journal of Multilingual and Multicultural Development*, *25*, 204-222. doi: 10.1080/01434630408666529

Dewaele, J.-M. (2010). 'Christ fucking shit merde!' Language preferences for swearing among maximally proficient multilinguals. *Sociolinguistic Studies*, *4*, 595-614. doi:10.1558/sols.v4i3.595

de Langhe, B., Puntoni, S., Fernandes, D., & van Osselaer, S. M. J. (2011). The anchor contraction effect in international marketing research. *Journal of Marketing Research*, *48*, 366-380. doi: http://dx.doi.org/10.1509/jmkr.48.2.366

Duke, A. A,& Bègue, L. (2015). The drunk utilitarian: Blood alcohol concentration predicts

utilitarian responses in moral dilemmas. *Cognition*, *134*,121-127.

doi:10.1016/j.cognition.2014.09.006

Eilola, T. M., Havelka, J., & Sharma, D. (2007). Emotional activation in the first and second

language. *Cognition and Emotion*, *21*, 1064-1076. doi:10.1080/02699930601054109

Erickson, T. D., & Mattson, M. E. (1981). From words to meaning: A semantic illusion.

Journal of Verbal Learning and Verbal Behavior, *20*, 540-551. doi:10.1016/S0022-

5371(81)90165-1

Eyal, T., Liberman, N., & Trope, Y. (2008). Judging near and distant virtue and vice. *Journal

of Experimental Social Psychology*, *44*, 1204-1209. doi:10.1016/j.jesp.2008.03.012

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical

power analysis program for the social, behavioral, and biomedical sciences. *Behavior

Research Methods*, *39*, 175-191. doi: 10.3758/BF03193146

Foot, P. (1978). *Virtues and vices and other essays in moral philosophy*. Berkeley: University

of California Press.

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic

Perspectives*, *19*(4), 25-42. doi: 10.1257/089533005775196732

Gawinkowska, M., Paradowski, M. B., & Bilewic, M. (2013). Second language as an

exemptor from sociocultural norms. Emotion-related language choice revisited. *PLoS

ONE*, *8*, e81225. doi:10.1371/journal.pone.0081225

Geipel, J., Hadjichristidis, C., & Surian, L. (2014, May). *Moral judgments: Foreign language

promotes utilitarian judgments*. Poster session presented at the 26th Convention of the

Association for Psychological Science, San Francisco, CA.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets

of moral foundations. *Journal of Personality And Social Psychology*, *96*, 1029-1046.

doi:10.1037/a0015141

Greene, J. (2013). *Moral tribes: Emotion, reason, and the gap between us and them*. New

York, NY, US: Penguin Press.

Greene, Joshua D. (in press). Beyond point-and-shoot morality: Why cognitive

(neuro)science matters for ethics. *Ethics*, *124*, 695-726. Retrieved from

http://www.wjh.harvard.edu/~jgreene/GreeneWJH/Greene-Ethics-PrePub-1-14-14.pdf

Greene, J. D., Sommerville, R., 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

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

Guerra, V. M., & Giner-Sorolla, R. (2010). The Community, Autonomy, and Divinity Scale

(CADS): A new tool for the cross-cultural study of morality. *Journal of Cross-Cultural

Psychology*, *41*, 35-50. doi:10.1177/0022022109348919

Hadjichristidis, C., Geipel, J., & Savadori, L. (in press). The effect of foreign language in

judgments of risk and benefit: The role of affect. *Journal of Experimental Psychology:

Applied.*

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to

moral judgment. *Psychological Review*, *108*, 814-834. doi:10.1037/0033-

295X.108.4.814

Haidt, J., & Joseph, C. (2008). The moral mind: How five sets of innate intuitions guide the

development of many culture-specific virtues, and perhaps even modules. In P.

Carruthers, S. Laurence, S. Stich (Eds.), *The innate mind Volume 3: Foundations and the future* (pp. 367-391). New York, NY, US: Oxford University Press. doi:10.1093/acprof:oso/9780195332834.003.0019

Haidt, J., Koller, S., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology*, *65*, 613-628. doi:10.1037/0022-3514.65.4.613

Hare, R. M. (1981). *Moral thinking: Its levels, method, and point*. Oxford: Oxford University Press.

Harris, C. L., Ayçiçeĝi, A., & Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. *Applied Psycholinguistics*, *24*, 561-579. doi:10.1017/S0142716403000286

Harris, C. L., Gleason, J. B., & Ayçiçeği, A. (2006). When is a first language more emotional? Psychophysiological evidence from bilingual speakers. In A. Pavlenko (Ed.), *Bilingual minds: Emotional experience, expression, and representation.* Clevedon, United Kingdom: Multilingual Matters.

Kahane, G., Everett, J. A.C., Earp, B. D.,Farias, M., & Savulescu, J. (2015). 'Utilitarian' judgments in sacrificial moral dilemmas do not reflect impartial concern for the greater good. *Cognition*, *134*, 193–209. doi: 10.1016/j.cognition.2014.10.005

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49-81). New York, NY, US: Cambridge University Press.

Keysar, B., Hayakawa, S. L., & An, S. (2012). The foreign-language effect: Thinking in a foreign tongue reduces decision biases. *Psychological Science*, *23*, 661-668. doi:10.1177/0956797611432178

Khemiri, L., Guterstam, J., Franck, J., & Jayaram-Lindström, N. (2012). Alcohol dependence

associated with increased utilitarian moral judgment: A case control study. *PLoS*

ONE, *7*, e39882. doi:10.1371/journal.pone.0039882

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., & Damasio, A.

(2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature*,

446, 908-911. doi: 10.1038/nature06785

Lucas, B. J., & Livingston, R. W. (2014). Feeling socially connected increases utilitarian

choices in moral dilemmas. *Journal of Experimental Social Psychology*, *53*, 1-4.

doi:10.1016/j.jesp.2014.01.011

Marian, V., & Neisser, U. (2000). Language-dependent recall of autobiographical memories.

Journal of Experimental Psychology: General, *129*, 361-368. doi:10.1037/0096-

3445.129.3.361

Mata, A., Ferreira, M. B., & Sherman, S. J. (2013). The metacognitive advantage of

deliberative thinkers: A dual-process perspective on overconfidence. *Journal of*

Personality and Social Psychology, *105*, 353-373. doi:10.1037/a0033640

Mendez, M. F., Anderson, E., & Shapira, J. S. (2005). An investigation of moral judgement

in frontotemporal dementia. *Cognitive and Behavioral Neurology*, *18*, 193-197.

doi:10.1097/01.wnn.0000191292.17964.bb

McFarland, S., Webb, M., & Brown, D. (2012). All humanity is my ingroup: A measure and

studies of identification with all humanity. *Journal of Personality and Social*

Psychology, *103*, 830-853. doi:10.1037/a0028724

Pavlenko, A. (2004). "Stop doing that, "la komu skazala!"": Language choice and emotions

in parent-child communication. *Journal of Multilingual and Multicultural*

Development, *25*, 179-203. doi: 10.1080/0143463040866528

Pavlenko, A. (2008). Emotion and emotion-laden words in the bilingual lexicon.

   Bilingualism: Language and Cognition, *11*, 147-164. doi:

   10.1017/S1366728908003283

Pavlenko, A. (2012). Affective processing in bilingual speakers: Disembodied cognition?

   International Journal of Psychology, *47*, 405-428. doi:10.1080/00207594.2012.743665

Pellizzoni, S., Siegal, M., & Surian, L. (2010). Contact principle and utilitarian moral

   judgments in young children. *Developmental Science*, *13*, 265-270.

   doi: 10.1111/j.1467-7687.2009.00851.x

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect

   effects in simple mediation models. *Behavior Research Methods, Instruments &*

   Computers, *36*, 717-731. doi:10.3758/BF03206553

Reder, L. M., & Kusbit, G. W. (1991). Locus of the Moses Illusion: Imperfect Encoding,

   Retrieval or Match? *Journal of Memory and Language*, *30*, 385-406. doi:10.1016/0749-

   596X(91)90013-A

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition*, *49*, 11-36.

   doi: 10.1016/0010-0277(93)90034-S

Schrauf, R. W., & Rubin, D. C. (2000). Internal languages of retrieval: The bilingual

   encoding of memories for the personal past. *Memory & Cognition*, *28*, 616-623.

   doi:10.3758/BF03201251

Schrauf, R. W., & Rubin, D. C. (2004). The 'Language' and 'Feel' of Bilingual Memory:

   Mnemonic Traces. *Estudios De Sociolingüística: Linguas, Sociedades E Culturas*, *5*,

   21-39.

Shweder, R. A., Much, N. C., Mahapatra, M., & Park, L. (1997). The 'big three' of morality

   (autonomy, community, divinity) and the 'big three' explanations of suffering. In A. M.

Brandt, P. Rozin (Eds.), *Morality and health* (pp. 119-169). Florence, KY, US: Taylor & Frances/Routledge.

Slovic, P., Finucane, M., Peters, E., & MacGregor, D. G. (2002). The affect heuristic. In T. Gilovich, D. Griffin, D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 397-420). New York, NY, US: Cambridge University Press.

Song, H., & Schwarz, N. (2008). Fluency and the detection of misleading questions: Low processing fluency attenuates the Moses illusion. *Social Cognition*, *26*, 791-799. doi:10.1521/soco.2008.26.6.791

Sunstein, C. R. (2005). Moral heuristics. *Behavioral and Brain Sciences*, *28*, 531-573. doi:10.1017/S0140525X05000099

Sutton, T. M., Altarriba, J., Gianico, J. L., & Basnight-Brown, D. M. (2007). The automatic access of emotion: Emotional stroop effects in Spanish-English bilingual speakers. *Cognition and Emotion*, *21*, 1077-1090. doi:10.1080/02699930601054133

Thomson, J. (1985). The trolley problem. *Yale Law Journal*, *94,* 1395-1415. doi: 10.2307/796133

Tversky, A. & Kahneman, D. (1981). The Framing of decisions and the psychology of choice. *Science*, *211*, 453-458. doi: 10.1126/science.7455683

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063-1070. doi:10.1037/0022-3514.54.6.1063

Footnotes

[1] The principle of description invariance or extensionality holds that the way options are described should not influence a person's preferences about them.

[2] We follow Greene (2014) in using deontological and utilitarian to mean respectively "characteristically deontological" and "characteristically utilitarian" as a function of the response content, not the underlying motivation.

[3] The sample size was determined based on an a-priori power analysis using *G*power* (Faul, Erdfelder, Lang, & Buchner, 2007) with the following settings: statistical power = .80, effect size *f* = .35 (medium to large effect size, based on Geipel et al., 2014), *p* = .05, number of groups = 2 (language conditions), number of repeated measures = 4 (scenarios), correlation between repeated measures ρ = 0.3 (estimated). The analysis indicated a minimum sample size of 34. We tested more participants than the power analysis suggested because the present studies were conducted during classes in which a greater number of participants was available (this applies to all reported studies). In the present studies, no interim analyses or stopping rules were applied.

[4] To determine the appropriate sample size we conducted an a-priori power analysis utilizing the estimates from Study 1a: effect size *f* = .43, alpha level = .05, power = .8, and ρ = 0.4. The minimum sample size suitable to detect a main effect of language condition was 26. [5] We report 95% CIs unless otherwise stated.

[6] The sample size was determined via an a-priori power analysis using the estimates from Study 1b: effect size *f* = .29, alpha level = .05, power = .8, and ρ = 0.6. The indicated minimum sample size was 68.

[7] We thank Catherine Caldwell-Harris for suggesting this possibility to us.

[8] We thank the action editor, Roger Giner-Sorolla, and an anonymous reviewer for suggesting several of these alternative explanations.

[9] The appropriate sample size was calculated based on an a-priori sample size calculation using the estimates from Study 1b: effect size $f = .29$ (medium effect), alpha level = .05, power = .8, and $\rho = 0.6$. The analysis indicated a minimum sample size of 68.

Table 1

Scenarios Used in Studies 1, 2 and 3 (English Versions).

Scenario	Description
Moral	
Dog[1, 2, 3]	Frank's dog was killed by a car in front of his house. Frank had heard that in China people occasionally eat dog meat, and he was curious what it tasted like. So he cut up the body and cooked it and ate it for dinner. [Study 1a: Franz; Study 1b: Frank; Study 2: Franco; Study 3: Lorenzo]
Incest[1, 2, 3]	A brother and sister are alone in the house and decide to make love just once. The sister is already taking birth control pills and the brother uses a condom. They both enjoy the act but decide not to do it again. They promise each other to keep it a secret.
Exam[1, 2, 3]	A student who doesn't know the answers to some of the questions in an exam copies them from a student sitting in front of him. He doesn't get caught and he and the other student both get good grades. [Study 3: Silvia]
Flag[1, 2]	A woman is cleaning out her closet, and she finds a national flag. She decides to cut it up into small pieces and uses the pieces to clean the toilet.
Bonus[3]	Two employees have worked equally toward a project. The project went well so they are entitled to a collective bonus of 1000 Euros. The manager, Giulia, is a friend of one of the employees and wants to allocate the entire bonus to him. Giulia sends an email to the finance office, but the email never arrives due to a server failure. As a result, each employee gets 500 Euros.
Non-moral	
Train[3]	Francesca lives in Florence and would like to visit one of her friends in Imola. If she takes the Eurostar she has to pay 21 Euro but the service is very comfortable. If she instead takes the regional train she pays 6 Euro but the wagons are a bit cold and dirty. Francesca takes the regional train.
Brand[3]	Marco has a strong headache. He goes to the pharmacy with the intention of buying a headache medicine from Bayer. The pharmacy is out of the medicine from Bayer Marco was looking for, but has a generic product which is, in his words, "exactly the same" as the product he intended to buy. Marco buys the generic brand medicine.

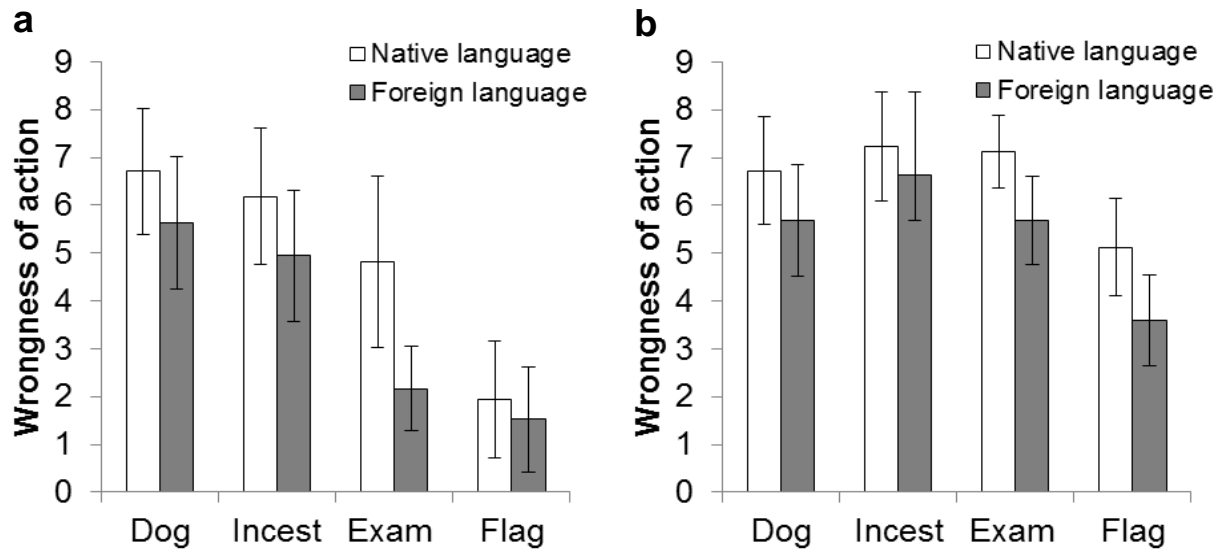*Note.* Superscripts indicate the studies in which the scenarios were used.

Figure 1. Mean wrongness of action ratings (0 = *perfectly ok*; 9 = *extremely wrong*) by scenario and

language condition. In Study 1a (a) the native language was German and the foreign language was

English. In Study 1b (b) the native language was Italian and the foreign language was English.
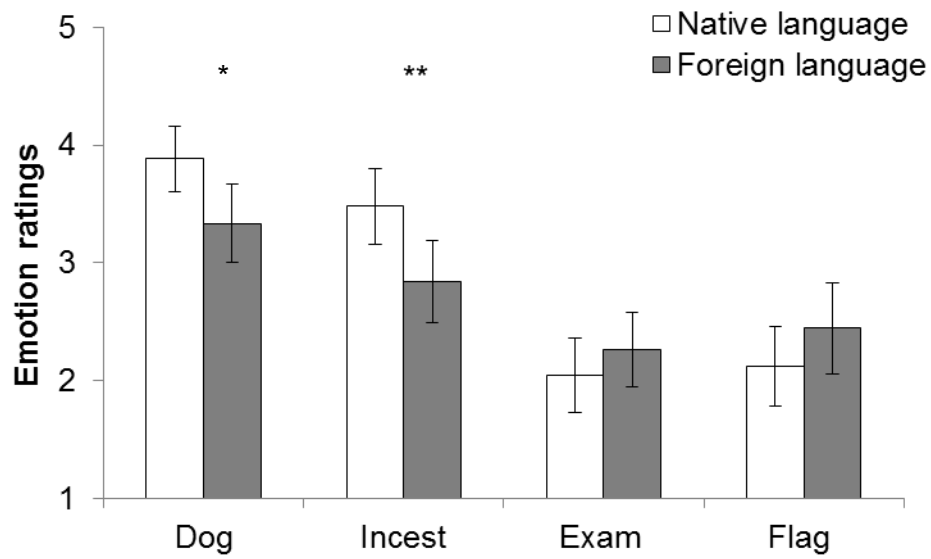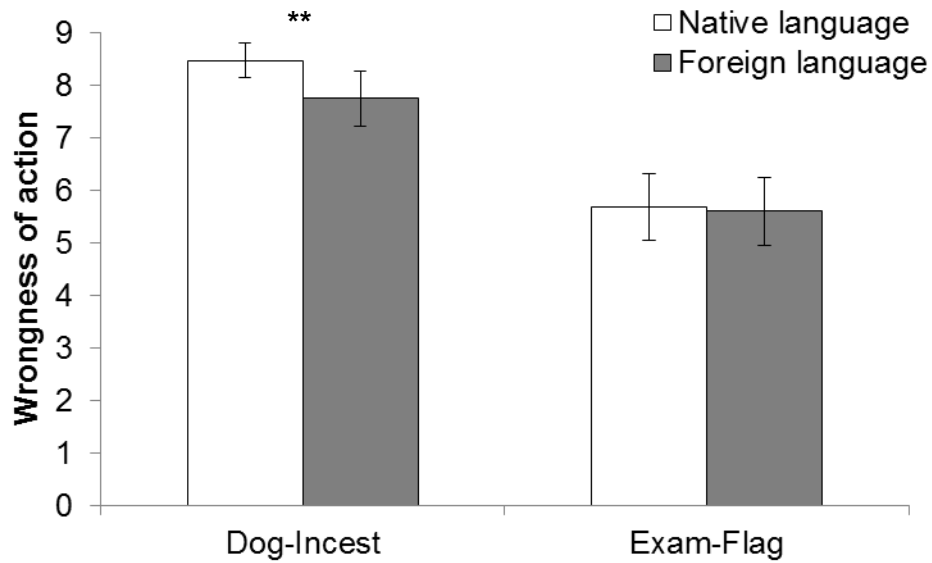
Error bars represent 95% CIs.

Figure 2. Mean emotion ratings by scenario and language condition (Study 2). Native language:

Italian; Foreign language: English. Error bars indicate 95% CIs. *$p < .05$,**$p < .01$.

Figure 3. Moral wrongness ratings for the groups of items by language condition (Study 2). Higher

scores indicate higher moral wrongness ratings. Native language: Italian; Foreign language:

English. Error bars represent 95% CIs. **$p < .01$.
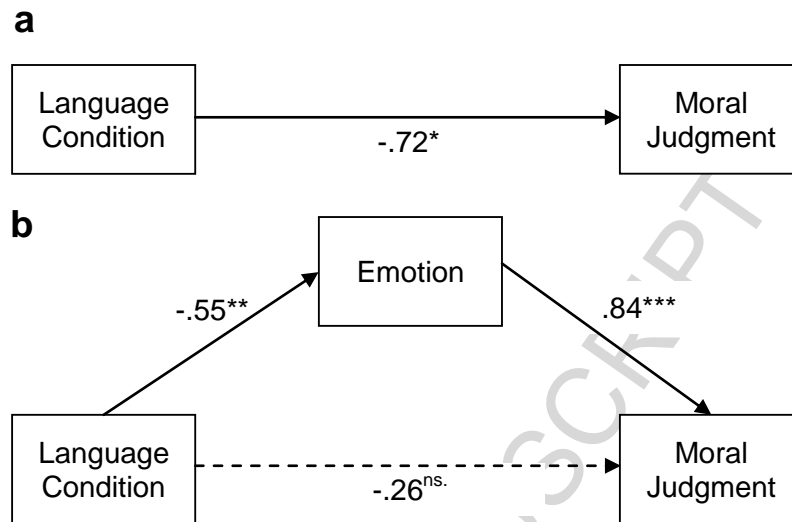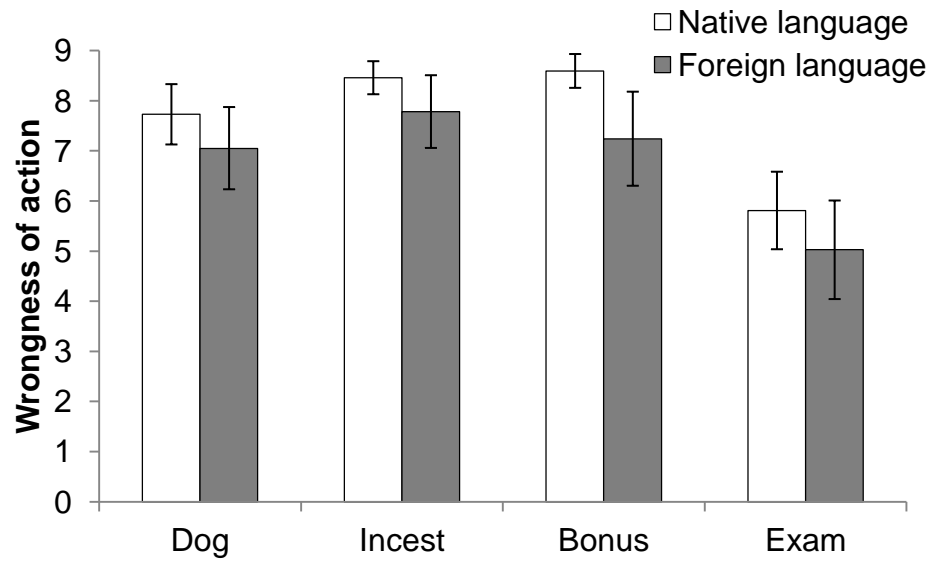
a

Figure 4. Illustration of the direct effect (a) and indirect effect (b) of language on moral judgment

(Study 2). Numbers refer to unstandardized beta weights. *$p < .05$ **$p < .01$ ***$p < .005$.

Figure 5. Moral wrongness ratings by scenario and language condition (Study 3). Native language:

Italian; Foreign language: German. Higher scores indicate higher moral wrongness ratings. Error

bars represent 95% CIs.
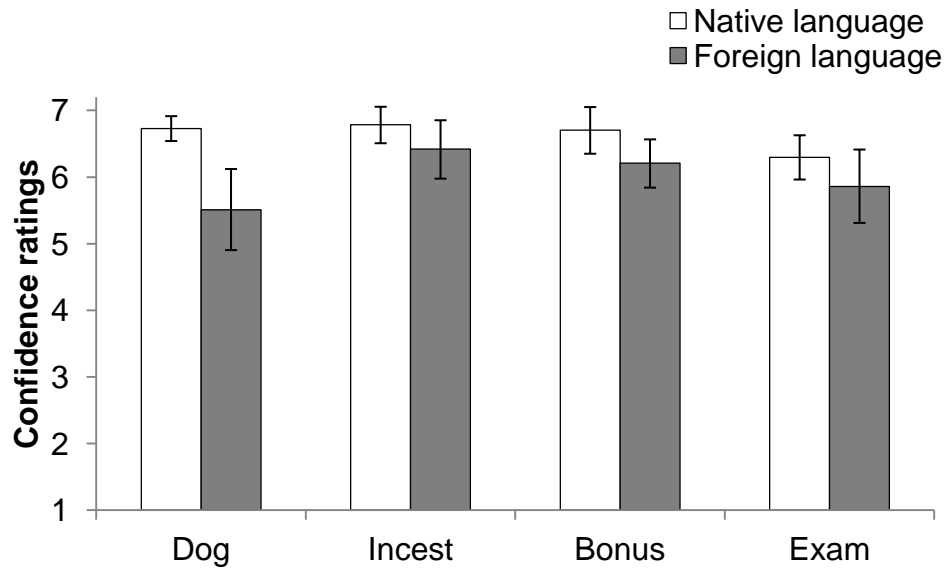
Figure 6. Confidence ratings in one's moral evaluations by scenario and language conditions (Study

3). Native language: Italian; Foreign language: German. Higher scores indicate higher confidence

ratings. Bars represent 95% CIs.

Appendix A

Details of Participants in the Foreign Language Conditions.

Table A.1

Details of participants in the foreign language conditions.

	Means, 95% CI
Study 1a (*n* = 19)	
Start age of	
English education	8.82, [8.13, 9.54]
Self-ratings of	
language skills in English[a]	3.82, [3.46, 4.14]
Study 1b (*n* = 35)[b]	
Start age of	
English education	9.29, [8.64, 9.91]
Self-ratings of	
language skills in English[a]	3.64, [3.45, 3.78]
Self-ratings of	
comprehension of the materials[c]	97%, [96%, 99%]
Study 2 (*n* = 42)[b]	
Start age of	
English education	8.40, [7.76, 8.98]
Self-ratings of	
language skills in English[a]	3.99, [3.87, 4.17]
Study 3 (*n* = 37)[b]	
Start age of	
German education	11.78, [10.42, 13.22]
Self-ratings of	
language skills in German[d]	3.72, [3.40, 4.00]

Note. [a]Participants evaluated their language skills in terms of conversational fluency, reading, writing, and understanding, each on a 5-point scale (1 = *almost none*, 2 = *poor*, 3 = *fair*, 4 = *good*, 5 = *very good;* scale adapted from Caldwell-Harris & Ayçiçeği-Dinn, 2009). Here we report the mean rating across these scales. [b]All participants had at least an intermediate level certificate (B1 or B2) in the foreign language as specified by the Common European Framework of Reference for Languages (CEFR; see page 24 in: http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf for descriptors). [c]We asked participants to rate how well they understood each scenario on a 6-point scale ranging from 50% (*some understanding*) to 100% (*excellent understanding*). [d]Participants evaluated their language skills in terms of reading and understanding, each on a 5-point scale (1 = *almost none*, 2 = *poor*, 3 = *fair*, 4 = *good*, 5 = *very good*). We report the mean rating across these scales.

Appendix B

English Versions of the Moses Illusion and Everyday Social and Moral Norms Tasks (Study 3)

Moses illusion task

This task was developed by Erickson and Mattson (1981) (see also Reder & Kusbit, 1991,

and Song & Schwarz, 2008). Following Song and Schwarz (2008), participants were instructed:

"You will read a couple of trivia questions and answer them. You can write the answer in
the blank. In case you do not know the answer, please write 'don't know.' You may or may
not encounter ill-formed questions which do not have correct answers if taken literally. For
instance, you might see the question 'Garfield is the dog of which cartoon?' In fact, Garfield
is not a dog it is a cat. Please, write 'can't say' for this type of question."

Following these instructions, participants had to respond to two questions:

(A) *"Which country is famous for cuckoo clocks, chocolate, banks, and pocket knives*?"
     (control question; correct answer: Switzerland)

(B) *"How many animals of each kind did Moses take on the Ark?"*
     (Moses illusion; correct answer: can't say)

Everyday social and moral norms

These materials were taken from Khemiri, Guterstam, Franck, and Jayaram-Lindström

(2012), who selected them from Mendez, Anderson, and Shapira (2005). Participants were given to

evaluate 15 items, each on a scale ranging from 1 (*not wrong*) to 4 (*severely wrong*).

In your opinion, how wrong is it to…

1) Fail to keep minor promises
2) Take the last seat on a crowded bus
3) Sell someone a defective car
4) Drive after having one drink
5) Cut in line when in a hurry
6) Don't give blood during blood drives
7) Are mean to someone you don't like
8) Say a white lie to get a reduced fare
9) Drive out the homeless from your neighborhood
10) Not help someone pick up their dropped papers
11) Keep excess-change at a store
12) Not offer to help after an accident
13) Ignore a hungry stranger
14) Fail to vote in minor elections
15) Keep money found on the ground

Highlights

- We investigated whether and how foreign language influences moral judgment.
- Foreign language prompted more lenient judgments for moral transgressions.
- Foreign language reduced confidence in people's moral evaluations.
- Violations of everyday norms were judged less harshly in a foreign language.
- Foreign language might act through a reduced activation of social and moral norms.