

BRIEF REPORT

Crime and Punishment: Morality Judgment in a Foreign Language

Evy Woumans, Ine Van der Cruyssen, and Wouter Duyck
Ghent University

The current study examined whether use of a foreign language affects the manner in which people evaluate a criminal situation. We employed a range of crime scenarios, for which severity judgment scores were obtained. Crimes that were written in a foreign language were systematically evaluated as less severe compared with the same cases described in the native language. We propose that these differences may be due to attenuated emotional processing in a nonnative language. Crucially, this observed variation in severity judgment may also affect magistrates and police interrogators confronted with crime scenarios formulated in a foreign tongue. This in turn would have inevitable consequences for the penalty they will or will not exact on the suspect.

Keywords: decision-making, crime judgment, foreign language effect, emotional attenuation, morality

Juries, judges, and the common person in the street are confronted with criminal acts on a daily basis, be it in the news, in films or series, or in real life. Sometimes accounts of these crimes are presented in people's native language, sometimes in a foreign language, for both laypeople and professionals. Imagine magistrates who are listening to a suspected felon describing the events of a supposed crime in another tongue known to the magistrates but not their native one. If the foreign language creates emotional detachment from the story or difficulty to vividly picture the occurrences, it may affect, and soften, their moral judgment.

Independent from the literature on crime judgment, previous research has indeed shown reduced emotional activation when

processing a foreign language (Pavlenko, 2005). For instance, the emotional weight of positive phrases such as *I love you* is much stronger in the native language (Dewaele, 2008), whereas taboo words, swear words, and childhood reprimands are experienced as less intense in the foreign language compared with the native one (Dewaele, 2004; Harris, 2004; Harris, Ayçiçeği, & Gleason, 2003). Moreover, negative words appear to be emotionally disembodied in foreign language speakers (Sheikh & Titone, 2016). Alternatively stated, the words are processed semantically but not, or to a lesser extent, affectively (Pavlenko, 2012). In addition, arousal felt as a result of lying is different when using a foreign language (Caldwell-Harris & Ayçiçeği-Dinn, 2009). The origin of these foreign language effects is generally assumed to lie in the context of language acquisition (Ivaz, Costa, & Duñabeitia, 2016; Schrauf, 2000). Whereas the native language is typically learned in a highly emotional setting (i.e., in attachment to parents and other caregivers), foreign languages are often acquired in formal situations (e.g., in a school context). Because humans' context-dependent memory is shaped by the language they use at the time of memory conception (Marian & Neisser, 2000; Thomson & Tulving, 1970), emotional concepts as well as norms and values are strongly associated with the mother tongue and may not be as easily accessible in the foreign language (see also Geipel, Hadjichristidis, & Surian, 2015).

This so-called foreign language effect has recently been applied to moral decision-making. It appears that individuals tend to sacrifice a person's life more readily in their foreign as opposed to their native language, a finding that has been dubbed the *moral foreign language effect*. In the commonly used footbridge dilemma, participants' willingness to push a large man onto a railway track to prevent five people from being killed by the trolley is gauged. There are two possible reactions to such a dilemma. If individuals sacrifice one life to save five, they are making a utilitarian decision, maximizing well-being for the majority of a population. If they do not, their response is deontological and they are operating under the principle that causing harm is morally

This article was published Online First January 13, 2020.

© Evy Woumans, Ine Van der Cruyssen, and Wouter Duyck, Department of Experimental Psychology, Ghent University.

Ine Van der Cruyssen is now at the Faculty of Social and Behavioural Sciences, University of Amsterdam.

This study was made possible by the Department of Experimental Psychology at Ghent University and is based on previous research conducted by Evy Woumans and Wouter Duyck, in which they aimed to determine whether use of a foreign language affects moral judgment. Wouter Duyck and Ine Van der Cruyssen developed the study concept together on the basis of Evy Woumans' prior work and contributed to the study design. Testing and data collection were performed by Ine Van der Cruyssen. Evy Woumans, and Ine Van der Cruyssen conducted data analyses. Evy Woumans drafted the article, and Wouter Duyck provided critical revisions. All authors approved the final version of the article for submission. Data and materials for this study can be accessed via Open Science Framework (<http://dx.doi.org/10.17605/OSF.IO/4YSQ2>). There was no prior dissemination of the ideas and data appearing in the article. All participants were healthy young adults. Their participation was entirely voluntary, and informed consent was obtained prior to the study.

Correspondence concerning this article should be addressed to Evy Woumans, Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, 9000 Gent, Belgium. E-mail: evy.woumans@ugent.be

wrong, regardless of the consequences. It is interesting that individuals more often push the man when this footbridge dilemma is presented in their foreign tongue (see Costa, Vives, & Corey, 2017 for a review). The dual-process theory of decision-making (Stanovich, 1999) posits that deontological acts (i.e., not pushing the man) are preferentially supported by an automated emotional response, whereas utilitarian action (i.e., pushing the man) is driven by cognitive (i.e., rational) reasoning.

The moral foreign language effect is therefore hypothesized to have either an emotional or cognitive locus. The former entails a decrease in automated emotional thinking after emotional attenuation resulting from foreign language use, whereas the latter supposes an increase in rational thinking and hence utilitarianism, precipitated by cognitive disfluency (e.g., Costa et al., 2014), which arises when employing a foreign language (Cop, Drieghe, & Duyck, 2015). However, a recent study on the validity of the dual-process theory showed that, in the majority of cases, utilitarian replies are already provided at the initial (automated) response stage, without deliberation having taken place (Bago & De Neys, 2019). This makes a cognitive account of the moral foreign language effect less plausible, because cognitive reasoning is not the main source of utilitarianism. In addition, the foreign language effect seems to arise from stunted regard for deontological rules rather than a heightened preference for utilitarianism (Hayakawa, Tannenbaum, Costa, Corey, & Keysar, 2017).

Emotional attenuation is thus a more likely candidate to explain reported foreign language effects in moral decision-making, but the picture may be somewhat more complicated and other factors may also play a role. For instance, Hayakawa and Keysar (2018) recently observed that participants report less vivid imagery of sensory experiences (e.g., viewing the sun as it is sinking below the horizon, smelling fresh paint) in their foreign language. In addition, these same authors demonstrated that this reduction in mental imagery partly accounts for the foreign language effect in moral decisions. It appears that visual imagery facilitates emotional processing, which in turn supports deontological action. This outcome is in line with previous research suggesting that visual interference blocks the ability to vividly picture a scene, hereby endorsing utilitarian acts, even in the mother tongue (Amit & Greene, 2012).

Nevertheless, the scenarios commonly employed in this type of decision-making research describe situations that are artificial and hardly realistic. Would a large man's body really be able to stop a railway trolley? And how likely is one to encounter such a situation? It is therefore difficult to draw conclusions about how use of a foreign tongue may affect decision-making in real life. In these unnatural circumstances, participants are hindered from relying on previous experiences and are thus prompted to fall back on moral heuristics. Heuristics are simple and efficient strategies derived from these previous experiences to help with an intricate problem at hand and reduce the cognitive load of making a decision. Still, it has been shown that so-called moral shortcuts or rules of thumb often lead to absurd moral judgments that have little relevance to real life (Sunstein, 2005). The present study therefore aimed to examine the foreign language effect for the first time regarding crime judgment, a socially relevant situation. Our investigation required having large samples of participants read murder scenarios in either their native or foreign language and judge the severity of the crimes. To facilitate comparison with previous research

employing personal perspective dilemmas, we also included personal conditions in which the agent of the otherwise third-person crime scene was the reader. Thus, similar to the case in narratives such as that about the footbridge, the reader would be the one causing harm and committing the crime. If the moral foreign language effect is indeed a case of blunted deontology (Hayakawa et al., 2017), we predict that crimes will be assessed as less severe in the foreign language.

Method

Participants

A total of 558 first bachelor's degree students in psychology at Ghent University in Belgium volunteered to participate in our study upon completing a consent form. An a priori G*Power analysis was conducted (Faul, Erdfelder, Buchner, & Lang, 2009) to calculate the sample size for an analysis of variance comparison of four groups (native vs. foreign language, and personal vs. third-person perspective) with an effect size of .25 and an actual power of .95. A total of 230 participants was required, which we exceeded. Because participation was entirely voluntary and the study contained only adult healthy participants, it was not necessary to obtain an institutional board review.

Participants were required to have only Dutch as their native and dominant language and English as their foreign language. English proficiency was reported for four language skills (comprehension, speaking, reading, writing) on a scale ranging from 1 (*not fluent*) to 7 (*very fluent*). A composite score was calculated to be used as a measure of general foreign language proficiency. Participants were so-called late learners of English who acquired it through formal education from their second year of secondary school (around the age of 12–13) onward. Some indicated earlier acquisition, through TV programs and computer games. Thirty-four individuals did not comply with the predefined language constraints (i.e., Dutch as the only native language) and were therefore excluded from the analysis. Another 12 individuals were removed because they reported being unable to sufficiently comprehend the content of one of the four scenarios. Comprehension was measured in terms of percentage of content understood per narrative, and 50% was used as a cutoff score. To be conservative, we decided to remove these individuals from all analyses because they often indicated poor comprehension of more than one scenario, and we inferred this was a reflection of their overall English reading skills. Background information and self-rated language proficiency of the remaining participants are reported in Table 1. Comprehension ratings can be found in Table 2.

Materials

Participants were asked to judge the severity of four different crime scenarios. These scenarios were obtained from the popular Belgian national tabloid *Humo* (Bulté, 2015a, 2015b) and were originally written in Dutch. Translations into English were made and were then back-translated into Dutch to ensure no loss of content or meaning had taken place (Brislin, 1970). Each scenario described a murder case for which severity was rated on a visual analogue scale ranging from 0 (*not severe*) to 9 (*one of the most severe crimes a person can commit*). Two of the four crimes were

Table 1
Demographic Information and Statistics on Comparisons by Language and Perspective Conditions

Variable	Personal perspective		Third-person perspective		χ^2	F	p
	Native language	Foreign language	Native language	Foreign language			
N	149	107	134	122			
Male/female ratio	18/131	19/88	18/116	15/107	$\chi^2(3, N = 512) = 2.035$	$F(3, 511) = 0.697$.565
Age in years: $M (SD)$	18.4 (1.8)	18.5 (2.4)	18.2 (0.8)	18.2 (1.1)			.554
Age of foreign language acquisition: $M (SD)$	11.9 (2.2)	11.8 (2.1)	12.2 (1.8)	11.9 (2.2)		$F(3, 511) = 0.758$.518
Foreign language proficiency ^a : $M (SD)$	5.5 (0.7)	5.3 (0.8)	5.4 (0.9)	5.5 (0.8)		$F(3, 511) = 1.379$.248

^a Measured on a scale from 1 (*not fluent*) to 7 (*very fluent*) for all skills. A composite score was calculated to use as a measure of general proficiency.

committed by men, whereas the two others were committed by women. In Scenario 1, a wife kills her husband in order to have extramarital sexual affairs but afterward regrets doing so. In Scenario 2, two sisters decide to murder their father, who had been physically and sexually abusing his children. Whereas one sister eventually confesses to the crime, the other one continues to deny it. Scenario 3 deals with a husband who has strangled and cut up his wife to come into her fortune, and Scenario 4 revolves around a rich man who weds a prostitute only to kill her when she gets pregnant and blame someone else for the murder. Regarding the content of the different scenarios, we aimed at variability. As such, the motives for murder varied greatly. Nevertheless, a small pilot study ($N = 7$) revealed no floor or ceiling effects in terms of severity ratings in any of the scenarios. The personal perspective adaptation involved substituting the agent from the third-person singular (e.g., the wife in Scenario 1) into the second-person singular (i.e., you). Because the scenarios were randomly assigned to participants, we decided not to change any gender-related pronouns. For example, female participants could receive a scenario that stated the man had killed his wife.

Procedure

Participants were randomly assigned to one of two language conditions (i.e., native vs. foreign) and one of two perspective conditions (i.e., personal vs. third-person). Given the large sample size in this homogeneous population, we employed a between-subjects design to avoid potential switch effects due to language or perspective change.

Each participant completed a pen-and-paper copy of the four scenarios in one of four conditions. The order of the four crimes was counterbalanced. They responded to each scenario by placing an X on the visual analogue scale ranging from 0 to 9. Each centimeter of the scale corresponded with an integer (i.e., 0–9) and each millimeter represented a decimal. As such, responses could be measured up to one digit after the separator (e.g., 3.2).

Results

Demographics

Analyses on demographic variables showed that groups created by randomly assigning participants to four possible conditions (Native vs. Foreign Language \times Personal vs. Third-Person Perspective) were matched in terms of male/female ratio, age, age of foreign language acquisition, and foreign language proficiency (see Table 1). Although scenario comprehension scores were high overall (see Table 2), it could be expected that they were slightly higher in the native as opposed to the foreign condition for each scenario (smallest $F = 17.058$, all $ps < .001$). There was also a difference between the two perspectives (smallest $F = 5.371$, highest $p = .021$), with higher scores in the third-person condition. No interaction was observed between the language and the perspective condition.

Scenarios

Because foreign language proficiency and scenario comprehension could influence severity judgment, we decided to take them

Table 2
Narrative Comprehension Ratings by Scenario for the Language and Perspective Conditions

Scenario	Personal perspective		Third-person perspective	
	Native language	Foreign language	Native language	Foreign language
Scenario 1 ^a	91.7 [89.9, 93.5]	88.3 [85.8, 90.9]	95.4 [94.0, 96.8]	90.9 [89.2, 92.7]
Scenario 2 ^b	92.9 [91.4, 94.4]	88.0 [85.7, 90.4]	95.1 [93.7, 96.7]	90.0 [88.1, 91.9]
Scenario 3 ^c	91.9 [90.2, 93.5]	87.7 [85.3, 90.2]	95.1 [93.6, 96.6]	89.0 [86.9, 91.0]
Scenario 4 ^d	92.7 [91.2, 94.3]	86.6 [84.0, 89.3]	94.4 [92.5, 96.2]	89.8 [87.6, 92.0]

Note. Comprehension was expressed in percentage understood, with 95% confidence intervals in brackets.

^a A wife kills her husband in order to have extramarital sexual affairs but afterward regrets doing so. ^b Two sisters murder their father, who had been physically and sexually abusing them. One sister eventually confesses to the crime, and the other one continues to deny it. ^c A husband strangles and cuts up his wife to come into her fortune. ^d A rich man weds a prostitute, kills her when she gets pregnant, and blames someone else for the murder.

up as covariates in our 2 (language: native vs. foreign) \times 2 (perspective: personal vs. third-person) analysis of covariance. All means are presented in Table 3. For Scenario 1, we determined a significant main effect of language, $F(1, 505) = 7.762, p = .006, \eta_p^2 = .015$, with lower severity ratings in the foreign language as opposed to the native condition. There was, however, no main effect of perspective, $F(1, 505) = 0.448, p = .504, \eta_p^2 = .001$, and no Language \times Perspective interaction, $F(1, 505) = 0.001, p = .981, \eta_p^2 < .001$. A similar effect of language was found for the other three scenarios: Scenario 2, $F(1, 505) = 17.977, p < .001, \eta_p^2 = .035$; Scenario 3, $F(1, 505) = 13.826, p < .001, \eta_p^2 = .027$; Scenario 4, $F(1, 507) = 6.136, p = .014, \eta_p^2 = .012$. There was never a main effect of perspective and also no Perspective \times Language interaction (highest F value = 1.602, all $ps > .206$). Regarding the covariates, scenario comprehension showed a positive relation to the dependent variable in Scenario 3, $F(1, 507) = 13.255, p < .001, \eta_p^2 = .026$, and Scenario 4, $F(1, 507) = 21.233, p < .001, \eta_p^2 = .041$, indicating that respondents tended to rate a scenario as less severe when comprehension was lower. Foreign language proficiency was not a significant covariate. Furthermore, no correlations were found between severity judgment and age of foreign language acquisition.

Discussion

The goal of the present study was to identify whether use of a foreign language influences crime judgment. Our results show that a crime is indeed deemed less severe when it is described in a language other than the native one. This was the case for all four murder scenarios employed in this study, making it the first demonstration ever that crime judgment is affected by language use. Hence, it is imperative that the outcome of our study be understood for the actual administration of justice in a globalized world, where suspects, juries, and judges often have different native languages than the official language of the jurisdiction in which they live.

Previous research on juridical judgment has already established that judges are prone to certain bias when it comes to crime assessment. As such, attractiveness and facial characteristics of the offender seem to determine the verdict and stringency of the sentence. Attractive individuals and individuals with large and round eyes, high eyebrows, and a narrow chin (i.e., baby-faced) tend to be judged less severely, an effect that has been observed in lab settings (e.g., Berry & Zebrowitz-McArthur, 1988; Efran, 1974; Leventhal & Kratoch, 1977) as well as in field studies (e.g., Stewart,

1980, 1985; Zebrowitz & McDonald, 1991). Although these findings are striking, personal appearance may be deemed a more plausible determinant of judgment than is the language employed. Intuitively, one could conceive that accounts of a crime are visualized in a juror's mind, creating a scenario independent of language. Nevertheless, similar findings have been reported within the literature on moral decision-making (Costa et al., 2017). Here, the prevailing theory ascribes the effect to emotional attenuation as a result of foreign language processing. Still, it may also be the case that use of a foreign language reduces the vividness of mental imagery (Hayakawa & Keysar, 2018), which in turn may hamper emotional processing. In any case, moral dilemmas involve personal harm for the respondents, because they are the ones deciding on matters of life and death. The current research now shows that the foreign language effect persists, even if a third person is the agent of the scenario and the respondent is merely a bystander.

Although the foreign language effect was observed systematically, our study is characterized by the fact that our crime cases varied in both motive and feelings of remorse. Furthermore, we noticed that narratives in which the agent was male were rated more gravely, by both male and female participants. It may be the case that crimes committed by men are judged as more appalling than are crimes committed by women. Conversely, it is possible that gender and severity coincidentally coincided. Notwithstanding, we believe that neither explanation alters the key finding of this study, namely the impact of foreign language use. Moreover, this may be regarded as a validation that the effect of language is persistent and does not depend on specific story characteristics. This being said, we must also point out that our participant sample consisted of both sexes but was still dominantly female (86%). Previous research has demonstrated that women may be more ethically sensitive than are men (e.g., Roxas & Stoneback, 2004). We must therefore note that the general outcome of this study could vary if the sample consists of only men, who may be less severe in their judgment. Nevertheless, there is again no indication that this may influence the effect of language reported here. Furthermore, because the materials in this study were transferred into the other language using the method of back-translation by a nonnative speaker, future work may need to determine whether materials written by native speakers would affect text perception, be it emotionally or otherwise.

All in all, our findings indicate that use of a foreign language diminishes crime severity judgment (see Figure 1), and although our study was conducted among laypeople, evidence from field studies into judgment bias related to the appearance of offenders suggests that our results may extrapolate to professionals. Crucially, the observed variation in morality standards may also affect magistrates and police interrogators confronted with crime scenarios formulated in a foreign tongue, who will consider these cases to be less severe than they would in their native one. This in turn would have inevitable consequences for the penalty they will or will not exact on the suspect. To illustrate, in the case of the European Court of Human Rights, the working language is either English or French. These are the languages in which judges deliberate, pleadings and written legal submissions are translated, and the judgment is drafted. Because this court often deals with important cases such a prohibition of torture, right to life, and right to liberty and security, it would have serious repercussions if judges were to underestimate the gravity of the situation merely

Table 3
Crime Severity by Scenario for the Language and Perspective Conditions

Scenario	Personal perspective		Third-person perspective	
	Native language	Foreign language	Native language	Foreign language
Scenario 1	7.3 [7.1, 7.5]	6.9 [6.5, 7.2]	7.4 [7.2, 7.6]	7.0 [6.7, 7.3]
Scenario 2	7.1 [6.9, 7.3]	6.3 [6.0, 6.6]	6.8 [6.5, 7.1]	6.3 [6.0, 6.6]
Scenario 3	8.0 [7.8, 8.1]	7.4 [7.1, 7.7]	8.0 [7.8, 8.2]	7.5 [7.2, 7.7]
Scenario 4	7.8 [7.6, 8.0]	7.3 [7.0, 7.6]	7.9 [7.7, 8.1]	7.5 [7.2, 7.7]

Note. Severity was rated on a scale from 0 (*not severe*) to 9 (*one of the most severe crimes a person can commit*), and 95% confidence intervals appear in brackets.

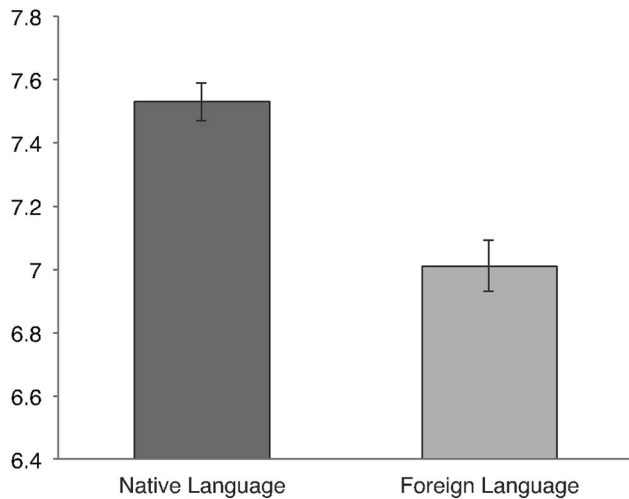


Figure 1. Mean crime severity judgment in native and foreign languages, collapsed over the four scenarios and indicated on a scale ranging from 0 (not severe) to 9 (one of the most severe crimes a person can commit). Error bars represent standard errors.

due to the language being used. Naturally, this is not the sole example. In an increasingly multilingual society (Grosjean, 2012), judges and juries all over the world may be operating in a nonnative language.

References

- Amit, E., & Greene, J. D. (2012). You see, the ends don't justify the means: Visual imagery and moral judgment. *Psychological Science*, 23, 861–868. <http://dx.doi.org/10.1177/0956797611434965>
- Bago, B., & De Neys, W. (2019). The intuitive greater good: Testing the corrective dual process model of moral cognition. *Journal of Experimental Psychology: General*, 148, 1782–1801. <http://dx.doi.org/10.1037/xge0000533>
- Berry, D. S., & Zebrowitz-McArthur, L. (1988). What's in a face? Facial maturity and the attribution of legal responsibility. *Personality and Social Psychology Bulletin*, 14, 23–33. <http://dx.doi.org/10.1177/0146167288141003>
- Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, 1, 185–216. <http://dx.doi.org/10.1177/135910457000100301>
- Bulté, A. (2015a). Dossier moord en doodslag in België. Deel 2: De grote verdwijning met het lijk [Murder and manslaughter file in Belgium: Part 2. The big disappearing trick with the corpse]. *Humo*, 3910, 22–27.
- Bulté, A. (2015b). Dossier moord en doodslag in België. Deel 3: Bijna perfecte moorden [Murder and manslaughter file in Belgium: Part 3. Almost perfect murders]. *Humo*, 3910, 36–40.
- Caldwell-Harris, C. L., & Ayçiçeği-Dinn, A. (2009). Emotion and lying in a non-native language. *International Journal of Psychophysiology*, 71, 193–204. <http://dx.doi.org/10.1016/j.ijpsycho.2008.09.006>
- Cop, U., Drieghe, D., & Duyck, W. (2015). Eye movement patterns in natural reading: A comparison of monolingual and bilingual reading of a novel. *PLoS ONE*, 10(8), e0134008. <http://dx.doi.org/10.1371/journal.pone.0134008>
- Costa, A., Foucart, A., Hayakawa, S., Aparici, M., Apesteguia, J., Heafner, J., & Keysar, B. (2014). Your morals depend on language. *PLoS ONE*, 9(4), e94842. <http://dx.doi.org/10.1371/journal.pone.0094842>
- Costa, A., Vives, M. L., & Corey, J. D. (2017). On language processing shaping decision making. *Current Directions in Psychological Science*, 26, 146–151. <http://dx.doi.org/10.1177/0963721416680263>
- 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. <http://dx.doi.org/10.1080/01434630408666529>
- Dewaele, J.-M. (2008). The emotional weight of *I love you* in multilinguals' languages. *Journal of Pragmatics*, 40, 1753–1780. <http://dx.doi.org/10.1016/j.pragma.2008.03.002>
- Efran, M. G. (1974). The effect of physical appearance on the judgment of guilt, interpersonal attraction, and severity of recommended punishment in a simulated jury task. *Journal of Research in Personality*, 8, 45–54. [http://dx.doi.org/10.1016/0092-6566\(74\)90044-0](http://dx.doi.org/10.1016/0092-6566(74)90044-0)
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149–1160. <http://dx.doi.org/10.3758/BRM.41.4.1149>
- Geipel, J., Hadjichristidis, C., & Surian, L. (2015). The foreign language effect on moral judgment: The role of emotions and norms. *PLoS ONE*, 10(7), e0131529. <http://dx.doi.org/10.1371/journal.pone.0131529>
- Grosjean, F. (2012). Bilingualism: A short introduction. In F. Grosjean & P. Li (Eds.), *The psycholinguistics of bilingualism* (pp. 5–25). Hoboken, NJ: Wiley.
- Harris, C. L. (2004). Bilingual speakers in the lab: Psychophysiological measures of emotional reactivity. *Journal of Multilingual and Multicultural Development*, 25, 223–247. <http://dx.doi.org/10.1080/01434630408666530>
- Harris, C. L., Ayçiçeği, 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. <http://dx.doi.org/10.1017/S0142716403000286>
- Hayakawa, S., & Keysar, B. (2018). Using a foreign language reduces mental imagery. *Cognition*, 173, 8–15. <http://dx.doi.org/10.1016/j.cognition.2017.12.010>
- Hayakawa, S., Tannenbaum, D., Costa, A., Corey, J. D., & Keysar, B. (2017). Thinking more or feeling less? Explaining the foreign-language effect on moral judgment. *Psychological Science*, 28, 1387–1397. <http://dx.doi.org/10.1177/0956797617720944>
- Ivaz, L., Costa, A., & Duñabeitia, J. A. (2016). The emotional impact of being myself: Emotions and foreign-language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42, 489–496. <http://dx.doi.org/10.1037/xlm0000179>
- Leventhal, G., & Kratoch, R. (1977). Physical attractiveness and severity sentencing. *Psychological Reports*, 40, 315–318. <http://dx.doi.org/10.2466/pr0.1977.40.1.315>
- Marian, V., & Neisser, U. (2000). Language-dependent recall of autobiographical memories. *Journal of Experimental Psychology: General*, 129, 361–368. <http://dx.doi.org/10.1037/0096-3445.129.3.361>
- Pavlenko, A. (2005). *Emotions and multilingualism*. New York, NY: Cambridge University Press.
- Pavlenko, A. (2012). Affective processing in bilingual speakers: Disembodied cognition? *International Journal of Psychology*, 47, 405–428. <http://dx.doi.org/10.1080/00207594.2012.743665>
- Roxas, M. L., & Stoneback, J. Y. (2004). The importance of gender across cultures in ethical decision-making. *Journal of Business Ethics*, 50, 149–165. <http://dx.doi.org/10.1023/B:BUSI.0000022127.51047.ef>
- Schrauf, R. W. (2000). Bilingual autobiographical memory: Experimental studies and clinical cases. *Culture & Psychology*, 6, 387–417. <http://dx.doi.org/10.1177/1354067X0064001>
- Sheikh, N. A., & Titone, D. (2016). The embodiment of emotional words in a second language: An eye-movement study. *Cognition and Emotion*, 30, 488–500. <http://dx.doi.org/10.1080/02699931.2015.1018144>
- Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. <http://dx.doi.org/10.4324/9781410603432>

- Stewart, J. E., II (1980). Defendant's attractiveness as a function in the outcome of criminal trials: An observational study. *Journal of Applied Psychology*, 10, 348–361.
- Stewart, J. E., II. (1985). Appearance and punishment: The attraction-leniency effect in the courtroom. *Journal of Social Psychology*, 125, 373–378. <http://dx.doi.org/10.1080/00224545.1985.9922900>
- Sunstein, C. R. (2005). Moral heuristics. *Behavioral and Brain Sciences*, 28, 531–542. <http://dx.doi.org/10.1017/S0140525X05000099>
- Thomson, D. M., & Tulving, E. (1970). Associative encoding and retrieval: Weak and strong cues. *Journal of Experimental Psychology*, 86, 255–262. <http://dx.doi.org/10.1037/h0029997>
- Zebrowitz, L. A., & McDonald, S. M. (1991). The impact of litigants' baby-facedness and attractiveness on adjudications in small claims courts. *Law and Human Behavior*, 15, 603–623. <http://dx.doi.org/10.1007/BF01065855>

Received August 22, 2019

Revision received November 25, 2019

Accepted December 12, 2019 ■