Decision-making depends on language: A meta-analysis of the Foreign Language Effect

Nicola Del Maschio1, Federico Crespi2, Francesca Peressotti3, Jubin Abutalebi1,4 and Simone Sulpizio2,5

1Centre for Neurolinguistics and Psycholinguistics (CNPL), Faculty of Psychology, Vita-Salute San Raffaele University, Milan, Italy; 2Department of Psychology, University of Milano-Bicocca, Milan Italy; 3Department of Developmental and Social Psychology, University of Padua, Padua, Italy; 4The Arctic University of Norway, Tromsø, Norway and 5Milan Center for Neuroscience (NeuroMi), University of Milano-Bicocca Milan, Italy

Abstract

In the present meta-analysis, we investigated the robustness and the magnitude of the Foreign Language Effect (FLE) – that is, the putative effect of language context (native versus foreign language) on decision-making. We also investigated whether the FLE is moderated by language experience – measured by second language age of acquisition and proficiency – or by methodological choices – the types of decision problems adopted, the presentation modality of the tasks administered, and the perspective in which problems are framed. Our results showed a reliable FLE, which was not moderated by language experience or methodological choices. We discuss our findings in relation to available theories of FLE, and indicate possible future directions to improve our understanding of the interplay between bilingualism and decision-making.

Introduction

As decision-makers in complex and volatile scenarios, we are constantly faced with the need to choose between alternative courses of action based on probabilistic cues and conflicting information. In addition to individual difference variables (e.g., demographic characteristics, cognitive ability, decision-making styles), psychological research has recently indicated that the language in which decisions are made is a contextual factor able to influence decision outcomes. In particular, systematically different choices have been reported when decision problems are presented in a native (L1) vs a foreign (L2) language. The systematic effect of language context on decision making has been termed Foreign Language Effect (FLE) (Keysar, Hayakawa & An, 2012).

Given that decisions are frequently presented to people in a second language in modern globalized societies, the implications of FLE for socio-economic and public health policies are obviously far-reaching. It has even been proposed that language could be used as a “nudge” to improve people’s decisions and guide interventions of policy makers (e.g., Costa, Vives & Corey, 2017). As appealing as this prospect may seem, however, the robustness, the magnitude, and the etiology of the FLE across extant research are yet to be tested. A meta-analysis on the current state of the literature would help setting the perimeter of the phenomenon, while exploring the potential contribution of moderator variables. Here, we present a comprehensive meta-analysis of studies on FLE on decision-making under conditions of risk and moral conflict.

A FLE was documented in decision-making involving risky prospects, where the decision maker cannot predict the outcome of a choice but knows the probabilities of all outcomes for alternative options (e.g., Hadjichristidis, Geipel & Savadori, 2015; Winskel, Ratitamkul, Brambley, Nagarachinda & Tiencharoen, 2016). One of the most popular paradigms in studies on the FLE on decision-making under risk is the Asian Disease Problem (Tversky & Kahneman, 1981). In this paradigm, the decision maker is confronted with a situation in which people can be saved from a pandemic disease by choosing between two alternative medicines. Although the consequences of the choice are identical in terms of number of saved lives, participants presented with one of two versions of the problem appear risk-seeking – i.e., they favor the riskier option – when the choice is framed in terms of losses (e.g., “400,000 out of 600,000 people will die”), and risk-averse – i.e., they favor the safer option – when the choice is framed in terms of gains (e.g., “200,000 out of 600,000 people will be saved”). This FRAMING EFFECT – the systematic tendency to make choices based on the form in which options are presented – has been shown to diminish when using an L2 (Costa, Foucart, Aparici & Aposteguía, 2014a; Keysar et al., 2012).

A FLE was documented also in moral dilemmas (e.g., Costa et al., 2014a; Geipel, Hadjichristidis & Surian, 2015a), where the decision maker can usually predict the outcome of a choice, knows the probabilities of all outcomes for alternative options, but is “pulled in
contrary directions by rival moral reasons” (Christensen & Gomila, 2012, p. 1251). For example, in the Footbridge Dilemma (Foot, 1978), an innocent bystander on a footbridge must be sacrificed to save five workers on a track from an out-of-control trolley moving in their direction. The decision maker, who must choose between sacrificing or not sacrificing the bystander, is confronted with two options: a) rejecting harm despite failing to maximize the number of saved lives, in accordance with the deontological perspective that the morality of actions is based on their intrinsic nature; b) maximizing the number of saved lives despite deliberately committing a harmful act, in accordance with the utilitarian perspective that the morality of actions is based on their outcomes. Whilst most respondents choose not to kill the bystander, utilitarian behavior has been reported to increase when the dilemma is presented in an L2 (e.g., Brouwer, 2019; Cipolletti, McFarlane & Weissglass, 2016; Geipel et al., 2015a).

Several explanatory hypotheses have been proposed to account for the FLE. In the following section, we synthesize the main ones.

**Etiology of the FLE**

**Enhanced cognitive control**

Dual-system theories in decision-making (e.g., Kahneman, 2003) and moral psychology (e.g., Greene, Morelli, Lowenberg, Nystrom & Cohen, 2008) propose that two systems are involved in decision-making processes: a fast, automatic, intuitive, and largely emotion-driven system (System 1), and a slower, systematic, deliberative and cognitively-controlled system that is also more effortful (System 2). According to Kahneman (2011), every contextual feature that increases mental stress or cognitive load, such as processing problems in a foreign language, could favor System 2 processes and/or reduce the influence of System 1 (but see Conway & Gawronski, 2013, who reported that cognitive load reduced participants’ utilitarian inclinations). Therefore, the FLE would be associated to a reduced reliance on System 1 and/or to an increased reliance on System 2. In particular, a foreign language context would promote the types of cognitive-controlled mechanisms that support more analytical appraisals and utilitarian decisions. The enhanced cognitive control hypothesis seems to suggest that the FLE will be beneficial to reasoning (or, at worst, neutral) in most circumstances. However, some results have disconfirmed this prediction (e.g., Geipel, Hadjichristidis & Surian, 2016), also showing that a foreign language context does not necessarily reduce cognitive biases when participants are presented with emotionally neutral tasks (e.g., Geipel et al., 2015a; Mækelæ & Pfuhl, 2019; Vives, Aparici & Costa, 2018). Moreover, using a process-dissociation approach, Hayakawa, Tannenbaum, Costa, Corey and Keysar (2017) reported that a foreign language reduced deontological inclinations and did not increase utilitarian tendencies. By interpreting these findings as a result of a dissociation process, the authors suggested that a foreign language may affect moral choice not through increased deliberation, but by dampening emotional reactions associated with the violation of deontological rules. Using the same process-dissociation approach, other studies reported that, when processing moral dilemmas in a foreign language, participants showed reduced levels of both utilitarian and deontological inclinations (Bialek, Paruzel-Czachura & Gawronski, 2019; Muda, Niszczoła, Bialek & Conway, 2018).

**Reduced emotionality**

Another hypothesis is that the FLE would depend on the reduced emotionality of decision-making contexts when these are framed in a foreign language. There is a two-step argument behind this hypothesis. On the one hand, emotions would promote intuitive, gut-feeling decisions that might cause biased reasoning (Greene, Nystrom, Engell, Darley & Cohen, 2004; Haidt, 2007). On the other hand, there is evidence for weakened emotional responses while processing an L2 (for reviews, see Caldwell-Harris, 2015; Pavlenko, 2017). On these grounds, the reduced emotionality account proposes that actively thinking in a foreign language would lead to decisions that are less distorted by emotional reactions (e.g., Keysar et al., 2012). The relative emotionality of a native vs a foreign language has been argued to be modulated by factors such as age of acquisition (AoA), language proficiency, language use and immersion, and (emotional) context of learning (Caldwell-Harris, 2015; Degner, Doycheva & Wentura, 2012; Dewealle, 2010; Sheikh & Titone, 2016). Recent findings also suggest that the extent to which the FLE occurs in proficient bilinguals can be influenced by the modality in which moral dilemmas are presented (auditory vs written) (Brouwer, 2019, 2020). Brouwer (2020), for example, reported an effect of foreign language when a sample of Dutch–English bilinguals listened to moral dilemmas in Dutch or English, but failed to report an effect when a different sample of Dutch–English bilinguals read the same dilemmas. One might speculate that the oral modality is the one through which language is learnt by children and most commonly used for day-to-day communication. Therefore, accessing semantic information through speech in L1 vs L2 may elicit a stronger emotional response in L1 than L2. Conversely, these differences may decrease when the same information is accessed and processed in a written format.

**Reduced access to social norms**

It was originally assumed that the framing effect of language would be visible only in case of emotionally-grounded biases (e.g., Costa et al., 2014a). However, optimal vs suboptimal decisions associated – respectively – with L2 and L1 may not necessarily depend on emotional distance. For example, emotion and FLE were recently dissociated in a study by Miozzo, Navarrete, Ongis, Mello, Girotto and Persott (2020), who showed that proficiently spoken Italian and Venetian (a regional language of Italy) elicited similar emotional responses, but yielded different decisions on both the Asian Disease Problem and the Footbridge Dilemma. As suggested by Geipel et al. (2015a, b), the discrepant decisions induced by native and foreign languages may also be due to a reduced accessibility of normative knowledge in foreign languages. It has been argued that individuals are usually exposed to normative knowledge early in life through social interactions mediated by their native language. Since episodic memories have been shown to include a trace of the language of encoding (e.g., Marian & Neisser, 2000; Schrauf & Rubin, 2000), a moral conflict presented in L1 may trigger greater language-dependent access to sociocultural and moral norms than a conflict presented in L2.

**Potential moderating factors of the FLE magnitude**

Foreign language effects were not reported ubiquitously in the literature. The inconsistency among previous studies may depend upon a number of factors. Here, we identified two groups of factors that could influence the magnitude of the FLE: 1) participants’ bilingual background; and 2) methodological design features.
Participants’ bilingual background

The FLE may be influenced by variability along the main quantifiable dimensions in which bilingual experience can be partitioned (i.e., L2 AoA, L2 proficiency, L2 exposure). Differences in bilinguals’ language background are known to affect bilingual language processing (for review, see Del Maschio & Abutaleb, 2019). It is therefore reasonable to hypothesize that differences in L2 experience will also modulate processing differences in decision-making when using the L1 vs an L2, with repercussions on the FLE. A role of L2 AoA in modulating language framing has been posited by emotion-based explanations of language effects (e.g., Costa et al., 2014a), as well as by accounts that hypothesize a reduced accessibility of normative knowledge through foreign languages (e.g., Geipel, Hadjichristidis & Surian, 2015b). Under both accounts, no FLE is expected when both languages are acquired early in similar contexts. On the one hand, a language learned in childhood should be especially emotional due to the emotional contexts of learning that are pervasive in childhood (see Caldwell-Harris, 2015). On the other hand, the occurrence of a FLE in e.g. a group of sequential bilinguals who learned their second language at school, compared to, say, a group of early or simultaneous bilinguals, may be attributable to a reduced access to moral and social norms when using the foreign language (Geipel et al., 2015a, 2015b). Another factor which has been proposed to modulate the FLE is L2 proficiency. Since emotional responses are expected to be equally intense with equally proficient languages (e.g., Anooshian & Hertel, 1994; Dewaele, 2004; Pavlenko, 2017), emotionally-grounded effects of language framing would tend to decrease in proficient L2 speakers. Alternatively, the cognitive load associated with processing problems in a foreign language would simply be reduced in fluent bilinguals, resulting in smaller differences in decision-making with a native and a foreign language (see Hayakawa et al., 2017). Nevertheless, the absence of the FLE in early and proficient bilinguals is debated (see Bialek & Fugelsang, 2019; Brouwer, 2020; Dylman & Champoux-Larsson, 2020; Miozzo et al., 2020). It is also possible that the relative frequency of daily use of L1 and L2 may affect the FLE. Even if the number of studies in the FLE literature that have specifically addressed this question is scarce, one may expect the FLE to decrease in bilingual individuals who are used to frequently dealing with decision problems in an L2. Conversely, situations in which the same problems are not faced with the same frequency in L1 and L2 should lead to a stronger FLE.

Methodological design features

In addition to the linguistic profile of bilingual participants, inconsistent results may also depend upon methodological design features of individual studies. Methodological choices include the types of problems adopted (i.e., decision problem vs moral dilemma), the presentation modality of the tasks administered to participants (i.e., auditory vs written), and the perspective in which problems are framed (i.e., personal vs impersonal). With regard to putative effects of problem type, facing problems under conditions of risk in L1 vs L2 may elicit different analytic strategies and/or different emotional responses than facing moral dilemmas in L1 vs L2. Different underlying mechanisms regulating decision-making under conditions of risk and moral conflict may be partially responsible for the inconsistency in detecting the FLE across previous studies. The occurrence and magnitude of the FLE may also depend on the modality in which problems and dilemmas are presented. As previously mentioned (see the Etiology of the FLE), the presentation modality of moral dilemmas has been shown to influence the occurrence of the FLE in proficient bilinguals (Brouwer, 2019, 2020). Another factor that appears to be important for the presence of the FLE is the distinction between personal and impersonal dilemmas. Under an emotion-based account of the language effect, Costa, Foucart, Hayakawa, Aparici, Apesteguia, Heafner, and Keysar (2014b) argued that a foreign language should induce more utilitarian decisions than a native language, and that the FLE may be stronger for personal vs impersonal dilemmas. However, some results have disconfirmed this prediction. For instance, this pattern was not replicated on the (impersonal) Switch dilemma (e.g., Cipolletti et al., 2016; Costa et al., 2014b). Moreover, Geipel et al. (2015b) presented multiple personal and impersomal dilemmas to participants, and reported that a FLE was present on some impersonal dilemmas and absent on some personal ones. The authors interpreted these findings as indicating that the FLE only occurs when dilemmas violate social or moral norms.

The present study

The primary objective of this meta-analysis is to assess the magnitude of the FLE by integrating behavioral evidence from both decision problems and moral dilemmas. To provide an estimate of the overall size of the FLE is critical, since previous findings about the occurrence and magnitude of the FLE are mixed. A further aim of this work is to examine whether and to what extent factors related to participants’ bilingual background (i.e., L2 AoA, L2 proficiency, L2 exposure) and methodological design features (i.e., Problem type, Task modality, Personal/Impersonal distinction) moderate meta-analytic results. Overall, establishing the boundaries and generalizability of FLE may possibly pave the way for a unitary account of the mechanisms underlying the effects of foreign languages on decision processing.

Materials and methods

Data collection and preparation

This meta-analysis is based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement guidelines (http://www.prisma-statement.org/; Liberati, Altman, Tetzlaff, Mulrow, Gøtzsche, Ioannidis, Clarke, Devereaux, Kleijnen & Moher, 2009; Moher, Liberati, Tetzlaff, Altman & Prisma Group, 2009). PRISMA’s goal is to improve the quality and the reliability of systematic reviews and meta-analyses by providing a set of common rules and recommendations for authors. PRISMA guidelines suggest to follow a 27-item checklist and report a flow diagram of the literature search and paper inclusion (Fig. 1). To identify all the available articles on the FLE, we performed an on-line literature search in three different databases – Scopus, Pubmed, and Web of Science –, which represent the main resources for psychological research. To be as inclusive as possible, the following input search keywords were used: “foreign language effect OR foreign-language effect”. Only studies written in English and published between 2012 and May 2020 were included (no article published before 2012 was found). This first search returned a total of 128 results. After this first step, we looked for any additional undetected article by inspecting the reference lists of the oldest articles (i.e., those published in 2012) or looking for peers’ recommendations, and 2 further articles were found.
From the initial set, we removed the duplicates, obtaining 65 results. A first screening based on title and abstract was independently conducted by two authors based on the following inclusion criteria: 1) empirical studies; 2) healthy adult bilinguals. Only peer-reviewed published journal articles were included. This first screening lead to a total of 50 eligible articles. These articles were then fully read to check whether they satisfied the inclusion criteria, which were extended to include the following: 3) testing the FLE; 4) using decision problems or moral dilemmas. Moreover, some exclusion criteria were also applied at this stage: 1) dependent variables different from the decision taken by participants (e.g., emotional response, degree of perceived morality); 2) presence of qualitative data/analyses only. We did not set any restriction on the specific languages spoken by each bilingual sample. The final sample included 15 articles (see Fig. 1), all multi-experiments studies. We assume that these studies were approved by their respective Ethics Committees prior to data collection.

**Data classification**

From each study, we aimed at extracting the following pieces of information: sample size, participants’ L1 and L2, L2 AoA, L2 proficiency, L2 exposure, Problem type (decision problem or moral dilemma – the problems and dilemmas of each study are reported in the Supplementary Materials), Task modality (auditory or written), Number of participants performing the task for each language (L1, L2), Number of participants choosing the emotional option for each language (i.e., choice due to a bias in decision problems and deontological choice in moral dilemmas), Number of participants opting, in each language, for the unbiased choice in decision problems and the utilitarian choice in moral dilemmas. These pieces of information are reported in Table 1.

In all experiments the manipulation was between-participants. L2 AoA was mostly measured by asking participants to self-report the onset age of L2 learning. In a few cases, it was defined by a label (e.g., childhood, late) or not reported. With regard to L2 proficiency, there was some heterogeneity in the way it was measured across studies. All studies except one reported subjective measures on a Likert scale. However, there was large variability in the highest value (from 5 to 30). To make the different scales comparable, each proficiency score was normalized using the following formula: \( x = \frac{(x-a)}{(b-a)} \), with \( x \) = the to-be-normalized score, \( a \) = the minimum value of the scale, \( b \) = the maximum value of the scale. Finally, most studies did not provide a quantitative measure of L2 exposure. A qualitative inference seems to suggest that participants had a low L2 exposure in approximately half of the included studies, whereas the exposure was high in the remaining studies.

**Data analysis**

Analyses were run with the *meta* package (Schwarzer, 2007) and the *dmetar* package (Harrer, Cuijpers, Furukawa & Ebert, 2019) in the R software (version 3.3.2). To investigate the FLE, we calculated the likelihood – expressed as Odd Ratio – to observe the emotional choice when problems were presented in L2 with respect to when they were presented in L1\(^1\). In other words, for moral dilemmas, we calculated the probability to make the deontological choice in L2 than in L1, whereas for decision problems we calculated the probability to make an emotional choice in L2 than in L1 when the problem is presented in the loss frame condition (e.g., for the *Asian disease problem*, this corresponds to the 400 people will die option, see Appendix for the full problem; note that this is also the result on which studies in the literature focus). We thus calculated L2 to L1 emotional choice odd ratios and corresponding two-tailed 95% confidence intervals for each experiment and then combined them to provide a pooled odds ratio and test for the overall effect (Z statistic). An odds ratio value greater than 1.0 indicates a higher tendency to opt for an unbiased choice in L2 than in L1, whereas a value equal to 1.0 corresponds to no difference between the two languages.

Because of the large heterogeneity among studies’ characteristics (e.g., the problems and dilemmas adopted in each study, in addition to differences related to the participants’ bilingual background), data were analyzed using the Hartung-Knapp-Sidik-Jonkman (HKSJ) mixed-effects model (IntHout, Ioannidis & Borm, 2014). To test for heterogeneity, the Q test and the I\(^2\) was considered. The Q test measures variability between the effect sizes in the sample of studies (Hedges & Olkin, 1985). The I\(^2\) is an index of the percentage of total variation (between effect sizes) across studies due to heterogeneity rather than chance (Higgins & Thompson, 2002). Values of 25%, 50%, and 75% have been described as indicating low, moderate and high heterogeneity, respectively (Higgins, Thompson, Deeks & Altman, 2003).

To investigate the presence of possible bias in the dataset, we examined the funnel plot and its asymmetry with the Egger’s t test (Egger, Smith, Schneider & Minder, 1997). The funnel plot represents the standard error (i.e., a measure of study precision) against the odd ratios of each individual study. In the absence of publication bias, one would expect the funnel plot to be symmetrical with studies symmetrically distributed around the center. An asymmetry, instead, would be suggestive of the presence of a possible bias. In this case, to evaluate the impact of the bias, Duval

\(^1\)Note that we opted for the emotional/deontological (instead of the rational/utilitarian) decision because it may be easier to interpret also by who is not familiar with the FLE literature.
Table 1. Characteristics of the studies included in the analysis.

| ID | Author       | Year | Journal             | Exp | Problem                    | Nr PP | Nr PP L1 E | Nr PP L2 E | Nr PP L1 UNB | Nr PP L2 UNB | Age   | Education       | L1     | L2     | AoA | Proficiency | Task Modality | Problem Type |
|----|--------------|------|---------------------|-----|---------------------------|-------|------------|------------|--------------|--------------|-------|----------------|--------|--------|-----|-------------|---------------|--------------|
| 1  | Costa et al. | 2014 | Cognition           | 1   | Asian disease problem     | 41    | 21         | 31         | 31           | 31           | 20.6  | University student | Spanish | English | 7.1 | 4.9 (7)     | Written       | DM            |
| 1  | Costa et al. | 2014 | Cognition           | 1   | Asian disease problem     | 40    | 30         | 27         | 37           | 37           | 20.1  | University student | Spanish | English | 6.2 | 6.2 (7)     | Written       | DM            |
| 1  | Costa et al. | 2014 | Cognition           | 2   | Discount                  | 31    | 40         | 37         | 43           | 43           | 21    | University student | Spanish | English | 9   | 4.75 (7)    | Written       | DM            |
| 1  | Costa et al. | 2014 | Cognition           | 2   | Ticket lost               | 64    | 7          | 48         | 22           | 22           | 21    | University student | Spanish | English | 8   | 4.75 (7)    | Written       | DM            |
| 2  | Miozzo et al.| 2020 | Cognition           | 2   | Asian disease problem     | 32    | 18         | 17         | 30           | 30.6         | NA    | Italian Venetian childhood | 9.2 (10) | Oral | DM            |
| 2  | Miozzo et al.| 2020 | Cognition           | 3   | Footbridge dilemma        | 68    | 43         | 86         | 28           | 41.5         | NA    | Italian Venetian childhood | 9 (10)   | Oral | MD            |
| 2  | Miozzo et al.| 2020 | Cognition           | 4   | Footbridge dilemma        | 347   | 61         | 316        | 136          | 36.6         | NA    | Italian Bergamasque | 8.1 (10) | Oral | MD            |
| 3  | Keysar et al.| 2012 | Psychol Sci         | 1a  | Asian disease problem     | 17    | 14         | 10         | 20           | 22           | 22    | University student | English | Japanese | 17 | 4.2 (7)     | Written       | DM            |
| 3  | Keysar et al.| 2012 | Psychol Sci         | 1b  | Asian disease problem     | 6     | 27         | 12         | 17           | 23           | 23    | University student | Korean  | English | 12 | 4.4 (10)    | Written       | DM            |
| 3  | Keysar et al.| 2012 | Psychol Sci         | 1c  | Asian disease problem     | 15    | 10         | 4          | 22           | 22           | 22    | University student | English | French  | 16 | 3.8 (10)    | Written       | DM            |
| 3  | Keysar et al.| 2012 | Psychol Sci         | 2   | Loss aversion (bets)      | 10    | 19         | 13         | 26           | 26           | 23    | University student | Korean  | English | 12 | 3.8 (10)    | Written       | DM            |
| 3  | Keysar et al.| 2012 | Psychol Sci         | 3   | Loss aversion (bets)      | 6     | 8          | 4          | 10           | 19           | 19    | University student | English | Spanish | 13 | 19 (30)     | Oral          | DM            |
| 4  | Brouwer      | 2019 | J Multiling Multicul| 1  | Crying baby               | 19    | 11         | 18         | 12           | 27.8         | 27.8  | University student | Dutch   | English | 19 | 3.5 (5)     | Oral          | MD            |
| 4  | Brouwer      | 2019 | J Multiling Multicul| 2  | Crying baby               | 7     | 23         | 9          | 21           | 27.8         | 27.8  | University student | Dutch   | English | 19 | 3.5 (5)     | Written       | MD            |
| 4  | Brouwer      | 2019 | J Multiling Multicul| 1  | Footbridge dilemma        | 27    | 3          | 19         | 11           | 11           | 27.8  | University student | Dutch   | English | 19 | 3.5 (5)     | Oral          | MD            |
| 4  | Brouwer      | 2019 | J Multiling Multicul| 2  | Footbridge dilemma        | 7     | 23         | 9          | 21           | 27.8         | 27.8  | University student | Dutch   | English | 19 | 3.5 (5)     | Written       | MD            |
| 4  | Brouwer      | 2019 | J Multiling Multicul| 1  | Lost wallet               | 26    | 4          | 19         | 11           | 11           | 27.8  | University student | Dutch   | English | 19 | 3.5 (5)     | Oral          | MD            |
| 4  | Brouwer      | 2019 | J Multiling Multicul| 2  | Lost wallet               | 30    | 0          | 24         | 6            | 27.8         | 27.8  | University student | Dutch   | English | 19 | 3.5 (5)     | Written       | MD            |
| 4  | Brouwer      | 2019 | J Multiling Multicul| 1  | Switch dilemma            | 14    | 16         | 9          | 21           | 29.5         | 29.5  | University student | Dutch   | English | 19 | 3.5 (5)     | Oral          | MD            |

(Continued)
| ID | Author               | Year | Journal | Exp | Problem                  | Nr PP L1 E | Nr PP L1 UNB | Nr PP L2 E | Nr PP L2 UNB | Age | Education               | L1 | L2 | AoA | Proficiency | Task Modality | Problem Type |
|----|----------------------|------|---------|-----|--------------------------|------------|--------------|------------|--------------|-----|-------------------------|----|----|-----|-------------|---------------|--------------|
| 4  | Brouwer              | 2019 | J Multiling Multicul | 2   | Switch dilemma           | 2          | 28           | 8          | 22           | 29.5| University student      | Dutch | English | 19  | 3.5 (5)     | Written | MD            |
| 4  | Brouwer              | 2019 | J Multiling Multicul | 1   | Taxes                    | 22         | 8            | 13         | 17           | 29.5| University student      | Dutch | English | 19  | 3.5 (5)     | Oral      | MD            |
| 4  | Brouwer              | 2019 | J Multiling Multicul | 2   | Taxes                    | 23         | 7            | 23         | 7            | 29.5| University student      | Dutch | English | 19  | 3.5 (5)     | Written | MD            |
| 4  | Brouwer              | 2019 | J Multiling Multicul | 1   | Vitamins                 | 16         | 14           | 18         | 12           | 29.5| University student      | Dutch | English | 19  | 3.5 (5)     | Oral      | MD            |
| 4  | Brouwer              | 2019 | J Multiling Multicul | 2   | vitamins                 | 16         | 14           | 12         | 18           | 29.5| University student      | Dutch | English | 19  | 3.5 (5)     | Written | MD            |
| 5  | Cavar & Tytus        | 2018 | J Multiling Multicul | 1   | Boat                     | 16         | 14           | 16         | 14           | 37.3| Non academic             | Croatian | German | 9.27 | 10.4 (12)  | Written | MD            |
| 5  | Cavar & Tytus        | 2018 | J Multiling Multicul | 1   | Footbridge dilemma      | 25         | 5            | 23         | 7            | 37.3| Non academic             | Croatian | German | 9.27 | 10.4 (12)  | Written | MD            |
| 5  | Cavar & Tytus        | 2018 | J Multiling Multicul | 1   | Hostage                  | 22         | 8            | 22         | 8            | 37.3| Non academic             | Croatian | German | 9.27 | 10.4 (12)  | Written | MD            |
| 5  | Cavar & Tytus        | 2018 | J Multiling Multicul | 1   | Soldier                  | 19         | 11           | 11         | 19           | 37.3| Non academic             | Croatian | German | 9.27 | 10.4 (12)  | Written | MD            |
| 5  | Cavar & Tytus        | 2018 | J Multiling Multicul | 1   | Submarine                | 15         | 15           | 17         | 13           | 37.3| Non academic             | Croatian | German | 9.27 | 10.4 (12)  | Written | MD            |
| 5  | Cavar & Tytus        | 2018 | J Multiling Multicul | 1   | Surgeon                  | 29         | 1            | 26         | 4            | 37.3| Non academic             | Croatian | German | 9.27 | 10.4 (12)  | Written | MD            |
| 6  | Winskel et al.       | 2016 | J Cogn Psychol      | 1   | Asian disease problem   | 32         | 18           | 25         | 27           | 19.8| University student      | Thai | English | 7.9  | 4.4 (10)    | Written | DM            |
| 6  | Winskel et al.       | 2016 | J Cogn Psychol      | 1   | Financial crisis        | 26         | 24           | 20         | 31           | 19.8| University student      | Thai | English | 7.7  | 4.5 (10)    | Written | DM            |
| 7  | Geipel et al.        | 2015 | Plos One            | 1a  | Footbridge dilemma      | 30         | 9            | 21         | 16           | 22.08| University student      | Italian | English | 10.7 | 3.3 (5)    | Written | MD            |
| 7  | Geipel et al.        | 2015 | Plos One            | 1b  | Footbridge dilemma      | 30         | 9            | 20         | 9            | 22.08| University student      | Italian | German | 13.3 | 3.8 (5)    | Written | MD            |
| 7  | Geipel et al.        | 2015 | Plos One            | 2   | Footbridge dilemma      | 56         | 6            | 77         | 22           | 22.08| University student      | Chinese | English | NA   | 3.1 (5)    | Written | MD            |
| 7  | Geipel et al.        | 2015 | Plos One            | 1a  | Trolley dilemma         | 18         | 21           | 10         | 27           | 23.41| University student      | Italian | English | 10.7 | 3.3 (5)    | Written | MD            |
| 7  | Geipel et al.        | 2015 | Plos One            | 1b  | Trolley dilemma         | 18         | 21           | 11         | 18           | 23.41| University student      | Italian | German | 13.3 | 3.8 (5)    | Written | MD            |
| 7  | Geipel et al.        | 2015 | Plos One            | 2   | Trolley dilemma         | 27         | 35           | 43         | 56           | 23.41| University student      | Chinese | English | NA   | 3.1 (5)    | Written | MD            |
| Study | Year | Journal | Task | University | Student Language | Task Language | Age | Gender | Condition | Method | Notes |
|-------|------|---------|------|------------|------------------|--------------|-----|--------|-----------|--------|-------|
| Hayakawa et al. | 2019 | Q J Exp Psychol | Taking a bet | Polish | English | | 15 | | Written | DM | |
| Hayakawa et al. | 2019 | Q J Exp Psychol | Taking a bet | Chinese | English | | 9 | | Written | DM | |
| Hayakawa et al. | 2019 | Q J Exp Psychol | Taking a bet | Spanish | English | | 12 | | Written | DM | |
| Corey et al. | 2017 | J Exp Psychol Learn | Button dilemma | University student | Spanish | English | 6.8 | | Written | MD | |
| Corey et al. | 2017 | J Exp Psychol Learn | Disable footbridge dilemma | University student | Spanish | English | 6.5 | | Written | MD | |
| Corey et al. | 2017 | J Exp Psychol Learn | Footbridge dilemma | University student | English | Spanish | 6.5 | | Written | MD | |
| Corey et al. | 2017 | J Exp Psychol Learn | Footbridge dilemma | University student | Spanish | English | 6.8 | | Written | MD | |
| Corey et al. | 2017 | J Exp Psychol Learn | Footbridge dilemma | University student | Spanish | English | 7.1 | | Written | MD | |
| Corey et al. | 2017 | J Exp Psychol Learn | Footbridge dilemma | University student | Spanish | English | 6.8 | | Written | MD | |
| Corey et al. | 2017 | J Exp Psychol Learn | Footbridge dilemma | University student | Spanish | English | 6.5 | | Written | MD | |
| Corey et al. | 2017 | J Exp Psychol Learn | Hospital dilemma | University student | Spanish | English | 6.1 | | Written | MD | |
| Corey et al. | 2017 | J Exp Psychol Learn | Injury footbridge dilemma | University student | Spanish | English | 6.8 | | Written | MD | |
| Corey et al. | 2017 | J Exp Psychol Learn | Switch dilemma | University student | English | Spanish | 6.5 | | Written | MD | |
| Corey et al. | 2017 | J Exp Psychol Learn | Switch dilemma | University student | Spanish | English | 7.1 | | Written | MD | |
| Corey et al. | 2017 | J Exp Psychol Learn | Switch dilemma | University student | Spanish | English | 6.8 | | Written | MD | |
| Corey et al. | 2017 | J Exp Psychol Learn | Switch dilemma | University student | Spanish | English | 6.5 | | Written | MD | |
| Corey et al. | 2017 | J Exp Psychol Learn | Switch dilemma | University student | Spanish | English | 6.5 | | Written | MD | |
| Corey et al. | 2017 | J Exp Psychol Learn | Switch dilemma | University student | Spanish | English | 6.8 | | Written | MD | |
| Corey et al. | 2017 | J Exp Psychol Learn | Terrorist dilemma | University student | Spanish | English | 6.1 | | Written | MD | |
| Cipolletti et al. | 2016 | Philos Psychol | Button dilemma | University student | Spanish | English | NA | | Written | MD | |

(Continued)
| ID | Author                          | Year | Journal      | Exp | Problem                      | Nr PP L1 | E | Nr PP L1 UNB | Nr PP L2 E | Nr PP L2 UNB | Age | Education         | L1 | L2 | AoA | Proficiency | Task Modality | Problem Type |
|----|--------------------------------|------|--------------|-----|------------------------------|----------|---|-------------|------------|-------------|-----|-------------------|----|----|-----|-------------|--------------|--------------|
| 10 | Cipolletti et al.              | 2016 | Philos Psychol | 2   | Footbridge dilemma           | 27       | 7 | 24          | 22         | 21.2        | University student | English | Spanish | NA | NA          | Written       | MD            |
| 10 | Cipolletti et al.              | 2016 | Philos Psychol | 3   | Switch dilemma               | 6        | 28| 9           | 37         | 21.2        | University student | English | Spanish | NA | NA          | Written       | MD            |
| 11 | Muda et al.                    | 2018 | J Exp Psychol Learn | 1  | Incongruent-congruent dilemma | 18      | 65| 32          | 51         | 20.9        | University student | Polish    | English | NA | > 5 (10)   | Written       | MD            |
| 12 | Brouwer                        | 2020 | Biling-Lang Cogn | 1   | Crying baby                  | 35      | 40| 36          | 43         | 25.5        | High education    | Dutch    | English | 9  | 4 (5)       | Oral          | MD            |
| 12 | Brouwer                        | 2020 | Biling-Lang Cogn | 2   | Crying baby                  | 49      | 26| 35          | 44         | 25.5        | High education    | Dutch    | English | 9  | 4 (5)       | Written       | MD            |
| 12 | Brouwer                        | 2020 | Biling-Lang Cogn | 1   | Footbridge dilemma           | 49      | 26| 22          | 57         | 25.5        | High education    | Dutch    | English | 9  | 4 (5)       | Oral          | MD            |
| 12 | Brouwer                        | 2020 | Biling-Lang Cogn | 2   | Footbridge dilemma           | 60      | 15| 65          | 14         | 25.5        | High education    | Dutch    | English | 9  | 4 (5)       | Written       | MD            |
| 12 | Brouwer                        | 2020 | Biling-Lang Cogn | 1   | Lost wallet                  | 33      | 42| 37          | 42         | 25.5        | High education    | Dutch    | English | 9  | 4 (5)       | Oral          | MD            |
| 12 | Brouwer                        | 2020 | Biling-Lang Cogn | 2   | Lost wallet                  | 6       | 69| 6           | 73         | 25.5        | High education    | Dutch    | English | 9  | 4 (5)       | Written       | MD            |
| 12 | Brouwer                        | 2020 | Biling-Lang Cogn | 1   | Switch dilemma               | 41      | 34| 43          | 36         | 25.5        | High education    | Dutch    | English | 9  | 4 (5)       | Oral          | MD            |
| 12 | Brouwer                        | 2020 | Biling-Lang Cogn | 2   | Switch dilemma               | 19      | 56| 26          | 53         | 25.5        | High education    | Dutch    | English | 9  | 4 (5)       | Written       | MD            |
| 12 | Brouwer                        | 2020 | Biling-Lang Cogn | 1   | Taxes                        | 28      | 47| 36          | 43         | 25.5        | High education    | Dutch    | English | 9  | 4 (5)       | Oral          | MD            |
| 12 | Brouwer                        | 2020 | Biling-Lang Cogn | 2   | Taxes                        | 64      | 11| 49          | 30         | 25.5        | High education    | Dutch    | English | 9  | 4 (5)       | Written       | MD            |
| 12 | Brouwer                        | 2020 | Biling-Lang Cogn | 1   | Vitamins                     | 32      | 43| 22          | 57         | 25.5        | High education    | Dutch    | English | 9  | 4 (5)       | Oral          | MD            |
| 12 | Brouwer                        | 2020 | Biling-Lang Cogn | 2   | Vitamins                     | 54      | 21| 48          | 31         | 25.5        | High education    | Dutch    | English | 9  | 4 (5)       | Written       | MD            |
| 13 | Costa et al.                   | 2014 | Plos One      | 1a  | footbridge dilemma           | 0       | 40| 8           | 32         | 21         | University student | Korean   | English | 14 | 2.85 (5)  | Written       | MD            |
| 13 | Costa et al.                   | 2014 | Plos One      | 1b  | Footbridge dilemma           | 17      | 44| 27          | 34         | 21         | University student | English | Spanish | 14 | 2.85 (5)  | Written       | MD            |
| 13 | Costa et al.                   | 2014 | Plos One      | 1c  | Footbridge dilemma           | 16      | 37| 18          | 36         | 21         | University student | English | French  | 14 | 2.85 (5)  | Written       | MD            |
| 13 | Costa et al.                   | 2014 | Plos One      | 1d  | Footbridge dilemma           | 2       | 16| 14          | 4          | 21         | University student | Spanish/ English | Hebrew  | 14 | 2.85 (5)  | Written       | MD            |
| ID | Author et al. | Year | Exp | Problem | Nr PP L1 E | Nr Pp L1 UNB | Nr PP L2 E | Nr PP L2 UNB | Age | Education | L1 | L2 | AoA | Proficiency | Task modality | Problem Type |
|----|--------------|------|-----|---------|------------|-------------|------------|------------|-----|-----------|----|----|-----|------------|---------------|-------------|
| 13 | Costa et al. | 2014 | Plos One | Footbridge dilemma | 32 | 148 | 79 | 101 | 21 | University student | Spanish/English | Spanish/English | 10.8 | 4.76(7) | Written | MD |
| 13 | Costa et al. | 2014 | Plos One | Switch dilemma | 147 | 36 | 146 | 36 | 21 | University student | Spanish/English | Spanish/English | 10.8 | 4.76(7) | Written | MD |
| 14 | Shin & Kim | 2017 | CrossMark | Switch dilemma | 38 | 43 | 31 | 49 | 20.2 | University student | Korean | English | 9.6 | 4.37(9) | Written | MD |
| 14 | Shin & Kim | 2017 | CrossMark | Fumes | 20 | 61 | 32 | 48 | 20.2 | University student | Korean | English | 9.6 | 4.37(9) | Written | MD |
| 14 | Shin & Kim | 2017 | CrossMark | Transplant | 80 | 1 | 58 | 22 | 20.2 | University student | Korean | English | 9.6 | 4.37(9) | Written | MD |
| 14 | Shin & Kim | 2017 | CrossMark | Crying baby | 46 | 35 | 28 | 62 | 20.2 | University student | Korean | English | 9.6 | 4.37(9) | Written | MD |
| 15 | Dylman & Champoux-Larsson | 2020 | Cognition | Asian disease problem | 122 | 47 | 117 | 45 | 27.6 | University student | Swedish | English | 10 | 7.5(10) | Written | DM |
| 15 | Dylman & Champoux-Larsson | 2020 | Cognition | Asian disease problem | 35 | 14 | 25 | 23 | 33.3 | University student | Swedish | French | late | 5.2(10) | Written | DM |
| 15 | Dylman & Champoux-Larsson | 2020 | Cognition | Footbridge dilemma | 88 | 14 | 82 | 14 | 32.8 | University student | Swedish | English | 10 | 7.6(10) | Written | MD |
| 15 | Dylman & Champoux-Larsson | 2020 | Cognition | Footbridge dilemma | 82 | 12 | 56 | 25 | 33.7 | University student | Swedish | French | 4.6(10) | Written | MD |
| 15 | Dylman & Champoux-Larsson | 2020 | Cognition | Footbridge dilemma | 130 | 28 | 116 | 31 | 33.6 | University student | Swedish | Norwegian | NA | 5.1(10) | Written | MD |
| 15 | Dylman & Champoux-Larsson | 2020 | Cognition | Footbridge dilemma | 130 | 30 | 113 | 22 | 33.7 | University student | Norwegian | Swedish | NA | 6.2(10) | Written | MD |

Note: ID = The number identifies experiments belonging to the same article; Author = first author’s surname; Year = year of article publication; Exp = number of the experiment as reported in the article; Problem = specific problem used in the experiment; Nr PP L1 E = number of participants in the L1 condition that chose the emotional option; Nr Pp L1 UNB = number of participants in the L1 condition that chose the unbiased option; Nr PP L2 E = number of participants in the L2 condition that chose the emotional option; Nr PP L2 UNB = number of participants in the L2 condition that chose the unbiased option; Age = participants’ mean age; Education = level of education as reported in the article; L1 = first language spoken by participants; L2 = second language spoken by participants; AoA = participants’ L2 AoA (as reported in the article); Proficiency = participants’ L2 (as reported in the article; the number in parentheses indicates the maximum value of the scale); Task modality = the modality in which the problems were presented; Problem Type = DM: Decision Making, MD = Moral Dilemma
and Tweedie’s (2000) trim-and-fill method was used. This method makes the funnel plot symmetrical by adding/omitting hypothetical studies to the plot when needed. In this way, the method provides an adjusted estimate of the effect size, which can be compared with the original one, thereby being informative on how much this effect is sensitive to the presence of (potential) missing studies. Note that although this method allows to investigate how sensitive the observed effect is to the presence of potential missing studies, it is not a way to calculate the actual values of the missing studies.

Finally, in order to test for the possible effects of inter-individual differences in L2 and methodological choices on the FLE, a series of meta-regressions were run which tested the effects of the following factors: L2 AoA, L2 proficiency, Problem type (decision problem vs moral dilemma), Task modality (auditory vs written) and the perspective in which problems are framed (personal vs impersonal).

**Results**

A total of 91 experiments were included in the analysis, totaling 13,886 participants – 6,928 tested in L1 and 6,958 tested in L2.

The qualitative analysis of the data summarized in Table 1 shows that most of the experiments (68%) involved participants speaking English as L2. The L1 distribution, instead, is more heterogeneous, with no language showing a large prevalence – the most used language is Dutch, which appears in 26% of the experiments. With regard to the type of task adopted, the majority of experiments (79%) used moral dilemmas. In these studies, the most used dilemma was the Footbridge dilemma, immediately followed by the Switch dilemma, which respectively occurred in 27% and 13% of the studies using dilemmas. The studies using decision problems (21% of the total) made large adoption of the Asian Disease Problem, which occurred in 47% of the studies using problems. Finally, with regard to the task modality, the vast majority of studies (82%) presented the task in a written format, with the auditory format being rarely adopted.

The forest plot in Figure 2 shows odds ratios for individual studies and the pooled odds ratio. The inspection of the last column (weight) indicates that all studies similarly contributed to the pooled effect, with weights ranging from a minimum of 0.3% (Costa et al., 2014b, Footbridge dilemma) to a maximum of 1.4% (e.g., Miozzo et al., 2020, Footbridge dilemma). Looking at the odds ratios for individual studies, the vast majority was in the expected direction.

The pooled odds ratio (OR) was significant (OR = 1.34, 95% CI [1.1 1.60], z = 3.16, p = .001). The result indicates that participants facing decision problems and moral dilemmas in their L2 are more likely to opt for unbiased or utilitarian decisions than participants facing problems and dilemmas in their L1. Put differently, a native and a foreign language would produce unbiased judgments than participants facing problems or dilemmas in their L2. The result indicates that the effect is to be expected for 95% of similar future studies. In our study, the prediction interval is [0.29 6.11]. Other than being wider than the CI, it also includes values smaller than 1. This indicates that a new similar study on FLE might also find a different result from that of our meta-analysis, i.e., either no effect of the language on the rational-vs-emotional option selection, or even a less likely unbiased option for problems presented in L2 than in L1. However, most of the prediction interval develops above the value of 1, suggesting that the great majority of similar future studies should be able to find the FLE.

**Discussion**

The present meta-analysis investigated the effects of foreign languages on decision-making under conditions of risk and moral conflict. Given the upsurge of new studies and the increasing interest within research in the psychology of judgment and decision-making, we aimed to quantify the reliability and strength of previous findings by synthesizing prior research. In addition, we were able to examine whether the overall size of the FLE was moderated by variables related to bilingual language experience (i.e., L2 AoA, L2 proficiency) and methodological design features (i.e., Problem type, Task modality, Personal/impersonal distinction).

Results revealed a small but significant FLE, indicating that participants facing problems and dilemmas in their L2 are more likely to produce unbiased judgments than participants facing problems or dilemmas in their L1. Put differently, a native and a foreign language would differentially impact heuristics and biases, with the latter being less conductive to framing effects and costly mistakes. As with any meta-analysis, the possibility exists that, if the literature synthesized is affected by some selection bias (e.g., publication bias), the conclusions drawn from published research might be overstated (see Rothstein, Sutton & Borenstein, 2005). Studies with medium-to-large standard error showing no effect or an opposite effect are completely missing. Moreover, the studies seem to be non-symmetrically distributed, showing a right prevalence. To evaluate the possible presence of publication bias, we ran the Egger’s test. The result was not significant (t < 1, p >.3), indicating no evidence for a publication bias.

As a last step, we tested the effect of possible factors in moderating the FLE. None of them showed a significant result (L2 AoA: beta = 0.01, SE = 0.1, z = 1.03, p > .3; L2 proficiency: beta = -0.69, SE = 0.88, z < 1, p >.4; Problem type: beta = -0.28, SE = 0.22, z = -1.23, p >.2, and Task modality: beta = 0.23, SE = 0.24, z < 1, p >.3; Personal/impersonal distinction: beta = 0.19, SE = 0.18, z = 1.03, p >.3).
modulating language framing (e.g., Costa et al., 2014a; Geipel et al., 2015b). In previous research, the FLE has been tested more frequently and reported more often in a specific group of bilinguals—that is, sequential bilinguals who learned a second language in school. However, a few studies found a language effect in simultaneous bilinguals (Miozzo et al., 2020) or no interaction between language effect and L2 AoA (e.g., Hayakawa, Lau, Holtzmann, Costa & Keysar, 2019). If the FL is learned in a classroom context, it is likely that the emotional connotation that characterizes the vocabulary of the L2 is not as rich as that of the vocabulary of L1, which is used in less formal, daily-life interactions (Caldwell-Harris, 2015; Costa et al., 2014a; see also Devaele, 2004). Alternatively, a FLE in this specific group of bilinguals compared to, say, a group of simultaneous bilinguals, may also be attributable to a reduced access to norms when using the foreign language (Geipel et al., 2015a, 2015b). We found that the age of acquisition of the foreign language did not significantly moderate our meta-analytic results, thus failing to support (as well as to discriminate between) prior hypotheses on the effect of this variable on the FLE. It could simply be that we failed to detect an effect of AoA because some critical age groups were only marginally represented in our sampled data (e.g., 0–6; >19 years). However, one might also speculate that the 'age factor' per se—that is, the point at which L2 learning begins—is

Fig. 2. Forest plot showing the results of the meta-analysis. The pooled effect size with confidence interval is indicated at the bottom of the figure. E_L1 = number of emotional choices done in L1; E_L2 = number of emotional choices done in L2; OR = Odds Ratio; 95% CI = 95% Confidence Interval.
Our finding is also in line with results from a recent study on the potential effect of this variable on the FLE. We thus failed to support (as well as to discriminate between) the hypothesis that L2 proficiency on the FLE. Notwithstanding a sufficient variability in our sampled data, we did not find a moderating effect of L2 proficiency on the FLE. Harris, 2004; Pavlenko, 2017), emotionally-grounded effects of language framing would broadly decrease in proficient bilingual speakers. An alternative explanation is that the cognitive load associated with comprehending the foreign language text of a decision problem or a moral dilemma would be simply reduced in a fluent bilingual, levelling out differences in decision processing associated with L1 and L2 (see Hayakawa et al., 2017). Notwithstanding a sufficient variability in our sampled data, we did not find a moderating effect of L2 proficiency on the FLE. We thus failed to support (as well as to discriminate between) prior hypotheses on the potential effect of this variable on the FLE. This finding is also in line with results from a recent meta-analysis by Circi, Gatti, Russo and Vecchi (2021), where proficiency in the foreign language was not found to moderate the magnitude of the observed FLE. One explanation may be that very few primary studies employed objective measures of proficiency; while most relied on subjective self-reports, which also differed from each other in many ways, making comparisons across studies problematic. Since previous research has suggested that the relationship between self-reports and objective measures of proficiency tends to vary depending on the tests being used and the languages assessed (see e.g., Marian, Blumenfeld & Kaushanskaya, 2007), a recommendation for future studies is to assess language knowledge with objective measures and standardized instruments. Importantly, the assessment of language knowledge should not be restricted to L2, but be extended to individual differences in native language attainment. L1 proficiency may vary greatly between individuals (e.g., Dąbrowska, 2019), and this possibility is usually overlooked in our sample data. It is also worthwhile to mention that, like Circi and colleagues (2021), we were only able to evaluate the mean proficiency of each sample in each experiment, thus reducing the interindividual variability in this factor.

It was possible that the FLE would be also moderated by problem type (i.e., decision problem vs moral dilemma), task modality (i.e., auditory vs written) and the perspective in which problems are framed (i.e., personal vs. impersonal). Effects of problem type, task modality, and personal/impersonal distinction would help understanding the robustness of the FLE, revealing that it may be enhanced or inhibited as a function of risk vs moral conflict or environmental differences. In the present meta-analysis, neither a moderating effect of problem type, nor moderating effects of task modality and personal/impersonal distinction were detected. With respect to problem type, the lack of an effect might hint at the existence of common mechanisms regulating decision-making under conditions of risk and moral conflict. With regard to task modality, Brouwer (2019) failed to report an effect of foreign language when Dutch–English bilinguals read moral dilemmas in Dutch or English, but reported the effect when a different sample of Dutch–English bilinguals listened to the same dilemmas. Here, since the vast majority of sampled studies presented the task in a written format, the lack of a significant effect might simply depend on skewed data. As for the personal/ impersonal distinction, our data does not seem to support the reasoning that the FLE might be stronger for personal than impersonal dilemmas. The lack of a moderating effect may (cautiously) be interpreted as indicating that, rather than by an attenuation of emotions associated with a personal perspective, the FLE in moral decision-making is conditional upon the violation of social or moral norms (see e.g., Geipel et al., 2015b).

A limitation of the present meta-analysis may possibly be identified in the heterogeneity of the studies included. A high heterogeneity may stem from the large differences showed by the studies and experiments at multiple levels – e.g., the languages spoken by participants, the differences in the variables describing L2 experience, and the problems adopted. Of note, a high heterogeneity might in fact suggest the presence of some moderating variables. Although we failed to report any moderator effect, for some of them this might be due to the nature of the variables tested. For example, variables measuring L2 experience were suboptimally distributed, with some values being little represented (e.g., the vast majority of studies used late bilinguals with an intermediate level of proficiency). Moreover, for task modality, almost all studies used a written presentation format, making the two
levels of the factor largely unbalanced. These limitations should be overcome by future empirical investigations.

Finally, a possible route for future research regards a potential moderating effect of language distance on the occurrence or magnitude of the FLE. It is indeed possible that typological similarity between (a bilingual’s) two languages plays an additional role in shaping judgments in conditions of risk and moral conflict. A precondition to test this hypothesis, however, is to gather sufficient consensus towards a working metric of typological similarity between languages.

Conclusions

There are still several pending issues relating to the FLE that need to be tackled to push forward research and its potential applications. We reported a significant effect of FL across decision-making contexts involving risk and moral conflict. However, we failed to report any effect of the moderator variables that prior research allowed us to test, which did not help to challenge interpretations to previous FLE findings. In order to adjudicate between alternative accounts, future research will have to further investigate the potential impact of bilingual experience by carefully measuring participants’ L2 proficiency and exposure, as well as by considering more extreme groups such as simultaneous bilinguals. Also, alternative accounts may be ruled out by integrating behavioral evidence with physiological and neuroimaging data, as well as by extending FLE research to other decision-making contexts and classes of heuristics.

Supplementary Material. For supplementary material accompanying this paper, visit https://doi.org/10.1017/S1366728921001012

References

Anooshian LJ and Hertel PT (1994) Emotionality in free recall: Language specificity in bilingual memory. Cognition & Emotion 8, 503–514.

Bialek M and Fugelsang J (2019) No evidence for decreased foreign language effect in highly proficient and acculturated bilinguals: A commentary on Cavar and Tytus (2018). Journal of Multilingual and Multicultural Development 40, 679–686.

Bialek M, Paruzel-Czachura M and Gawronski B (2019) Foreign language effects on moral dilemma judgments: An analysis using the CNI model. Journal of Experimental Social Psychology 85, 103855.

Birdsong D (2018) Plasticity, variability and age in second language acquisition and bilingualism. Frontiers in psychology 9, 81.

Brouwer S (2019) The auditory foreign-language effect of moral decision making in highly proficient bilinguals. Journal of Multilingual and Multicultural Development 40, 865–878.

Brouwer S (2020) The interplay between emotion and modality in the foreign-language effect on moral decision making. Bilingualism: Language and Cognition, 1–8.

Caldwell-Harris CL (2015) Emotionality differences between a native and foreign language: Implications for everyday life. Current Directions in Psychological Science 24, 214–219.

Christensen JF and Gomila A (2012) Moral dilemmas in cognitive neuroscience: A principled review. Neuroscience & Biobehavioral Reviews 36, 1249–1264.

Cipolletti H, McFarlane S and Weissgall C (2016) The moral foreign-language effect. Philosophical Psychology 29, 23–40.

Ciri R, Gatti D, Russo V and Vecchi T (2021) The foreign language effect on decision-making: A meta-analysis. Psychonomic Bulletin & Review, 1–11.

Conway P and Gawronski B (2013) Deontological and utilitarian inclinations in moral decision making: a process dissociation approach. Journal of personality and social psychology 104, 216.

Corey JD, Hayakawa S, Foucart A, Aparici M, Botella J, Costa A and Keyser B (2017) Our moral choices are foreign to us. Journal of experimental psychology: Learning, Memory, and Cognition 43, 1109.

Costa A, Foucart A, Arnon I, Aparici M and Aposteguia J (2014a) “Piensa” twice: On the foreign language effect in decision making. Cognition 130, 236–254.

Costa A, Foucart A, Hayakawa S, Aparici M, Aposteguia J, Heefner J and Keyser B (2014b) Your morals depend on language. PLoS one 9, e94842.

Costa A, Vives ML and Corey JD (2017) On language processing shaping decision making. Current Directions in Psychological Science 26, 146–151.

Dąbrowska E (2019) Experience, aptitude, and individual differences in linguistic attainment: A comparison of native and nonnative speakers. Language Learning 69, 72–100.

Degner J, Doycheva C and Wentura D (2012) It matters how much you talk: On the automaticity of affirmative connotations of first and second language words. Bilingualism: Language and Cognition 15, 181–189.

Del Maschio N and Abutaleb J (2019) Language organization in the bilingual and multilingual brain. In: The handbook of the neuroscience of multilingualism. pp. 199–213.

Dewaele JM (2004) The emotional force of swearwords and taboo words in the speech of multilinguals. Journal of multilingual and multicultural development 25, 204–222.

Dewaele J (2010) Emotions in multiple languages. Palgrave Macmillan, London.

Duval S and Tweedie R (2000) Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 56, 455–463.

Dyman AS and Champoux-Larson ME (2020) It’s (not) all Greek to me: Boundaries of the foreign language effect. Cognition 196, 104148.

Egger M, Smith GD, Schneider M and Minder C (1997) Bias in Meta-Analysis Detected by a Simple, Graphical Test. BMJ 315, 629–34.

Foot P (1978) The problem of abortion and the doctrine of the double effect. Reprinted in Virtues and Vices and Other Essays in Moral Philosophy. Blackwell, Oxford, pp. 19–32.

Geipel J, Hadjichristidis C and Surian L (2015a) How foreign language shapes moral judgment. Journal of Experimental Social Psychology 59, 8–17.

Geipel J, Hadjichristidis C and Surian L (2015b) The foreign language effect on moral judgment: The role of emotions and norms. PLoS one 10, e0131529.r

Geipel J, Hadjichristidis C and Surian L (2016) Foreign language affects the contribution of intentions and outcomes to moral judgment. Cognition 154, 34–39.

Greene JD, Nystrom LE, Engell AD, Darley JM and Cohen JD (2004) The neural bases of cognitive conflict and control in moral judgment. Neuron 44, 389–400.

Greene JD, Morelli SA, Lowenberg K, Nystrom LE and Cohen JD (2008) Cognitive load selectively interferes with utilitarian moral judgment. Cognition 107, 1144–1154.

Hadjichristidis C, Geipel J and Savadori L (2015) The effect of foreign language in judgments of risk and benefit: The role of affect. Journal of Experimental Psychology: Applied 21, 117.

Haidt J (2007) The new synthesis in moral psychology. Science 316, 998–1002.

Harrer M, Cuijpers P, Furukawa T and Ebert DD (2019) dmetar: Companion R Package For The Guide ‘Doing Meta-Analysis in R’. R package version 0.0.9000. 2019. http://dmetar.protectlab.org.

Harris CL (2004) Bilingual speakers in the lab: Psychophysiological measures of emotional reactivity. Journal of multilingual and multicultural development 25, 223–247.

Hayakawa S, Tannenbaum D, Costa A, Corey JD and Keysar B (2017) Thinking more or feeling less? Explaining the foreign-language effect on moral judgment. Psychological science 28, 1387–1397.

Hayakawa S, Lau BKY, Holtzmann S, Costa A and Keysar B (2019) On the reliability of the foreign language effect on risk-taking. Quarterly Journal of Experimental Psychology 72, 29–40.

Hedges LV and Olkin I (1985) Statistical methods for meta-analysis. Orlando, FL: Academic Press.

Higgins JP and Thompson SG (2002) Quantifying heterogeneity in a meta-analysis. Statistics in medicine 21, 1539–1558.

Higgins JP, Thompson SG, Deeks JJ and Altman DG (2003) Measuring inconsistency in meta-analyses. BMJ 327, 557–560.
IntHout J, Ioannidis JP and Borm GF (2014) The Hartung-Knapp-Sidak-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC medical research methodology* 14, 25.

IntHout J, Ioannidis JP, Rovers MM and Goeman JJ (2016) Plea for routinely presenting prediction intervals in meta-analysis. *BMJ open* 6.

Kahneman D (2003) A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist* 58, 697–720.

Kahneman D (2011) *Thinking, fast and slow*. Farrar, Straus and Giroux.

Keysar B, Hayakawa SL and An SG (2012) The foreign-language effect: Thinking in a foreign tongue reduces decision biases. *Psychological science* 23, 661–668.

Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J and Moher D (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of clinical epidemiology* 62, e1-e34.

Mækelæ JM and Pfuhl G (2019) Deliberate reasoning is not affected by language. *PLOS ONE* 14, e0211428. https://doi.org/10.1371/journal.pone.0211428

Marian V and Neisser U (2000) Language-dependent recall of autobiographical memories. *Journal of Experimental Psychology: General* 129, 361.

Marian V, Blumenfeld HK and Kaushanskaya M (2007) The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research* 50, 940–967.

Miozzo M, Navarrete E, Ongis M, Mello E, Girotto V and Peressotti F (2020) Foreign language effect in decision-making: How foreign is it?. *Cognition* 199, 104245.

Moher D, Liberati A, Tetzlaff J, Altman DG, & Prisma Group. (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine* 6, e1000097.

Muda R, Niszczota P, Biały M and Conway P (2018) Reading dilemmas in a foreign language reduces both deontological and utilitarian response tendencies. *Journal of Experimental Psychology: Learning Memory and Cognition* 44, 321–326.

Pavlenko A (2017) Do you wish to waive your rights? Affect and decision-making in multilingual speakers. *Current opinion in Psychology* 17, 74–78.

Rothstein HR, Sutton AJ and Borenstein M (2005) Publication bias in meta-analysis. In Rothstein HR, Sutton AJ and Borenstein M (eds), *Publication Bias in Meta-Analysis Prevention, assessment and adjustments*. John Wiley & Sons, pp. 1–7.

Schrauf RW and Rubin DC (2000) Internal languages of retrieval: The bilingual encoding of memories for the personal past. *Memory & Cognition* 28, 616–623.

Sheikh NA and Titone D (2016) The embodiment of emotional words in a second language: An eye-movement study. *Cognition and Emotion* 30, 488–500.

Shin H and Kim J (2017) Foreign language effect and psychological distance. *Journal of psycholinguistic research* 46, 1339–1352.

Sterne JAC, Becker BJ and Egger M (2005) The funnel plot. In Rothstein H, Sutton A and Borenstein M (eds), *Publication bias in meta-analysis: Prevention, assessment, and adjustments*. West Sussex, England: John Wiley & Sons, pp. 75–98.

Tversky A and Kahneman D (1981) The framing of decisions and the psychology of choice. *Science* 211, 453–458.

Vives ML, Aparici M and Costa A (2018) The limits of the foreign language effect on decision-making: The case of the outcome bias and the representativeness heuristic. *PloS one* 13, e0203528.

Winskel H, Ratitamkul T, Brambley V, Nagarachinda T and Tiencaroon S (2016) Decision-making and the framing effect in a foreign and native language. *Journal of Cognitive Psychology* 28, 427–436.