

Exploring personality factors as modulators of the foreign language effect on moral decision-making

Barbara Braida¹, Javier Rodríguez-Ferreiro² and Mireia Hernández¹ 

Research Article

Cite this article: Braida, B., Rodríguez-Ferreiro, J. and Hernández, M. (2026). Exploring personality factors as modulators of the foreign language effect on moral decision-making. *Bilingualism: Language and Cognition* 1–15.
<https://doi.org/10.1017/S1366728926101217>

Received: 26 July 2024
Revised: 10 February 2026
Accepted: 20 February 2026

Keywords:

bilingualism; foreign language effect; moral decision-making; moral dilemmas; personality factors

Corresponding author:

Mireia Hernández;
Email: mireiahernandez@ub.edu

¹Cognition and Brain Plasticity Unit (CBPU), Departament de Cognició, Desenvolupament i Psicologia de l'Educació, Secció Processos Cognitius, Institut de Neurociències (UBneuro), Universitat de Barcelona (UB), Barcelona, Spain and
²Grup de Recerca en Cognició i Llenguatge (GRECIL), Departament de Cognició, Desenvolupament i Psicologia de l'Educació, Secció Processos Cognitius, Institut de Neurociències (UBneuro), Universitat de Barcelona, Barcelona, Spain

Abstract

This study investigates how personality traits modulate the foreign language effect (FLe) in moral decision-making. Meta-analyses have shown that the FLe has a small-to-moderate effect size, suggesting that individual characteristics may constrain its impact. We hypothesized that traits promoting strong deontological tendencies would moderate this effect. Our results showed that lower conscientiousness and higher emotional stability were associated with more resilient deontological responses, showing reduced susceptibility to the FLe. However, deontological choices associated with lower extraversion were overridden in a foreign language. These findings offer the first empirical evidence that broad personality traits can modulate the FLe. Moreover, our results – statistically significant but modest in magnitude – highlight key methodological considerations, including dilemma types, trait measures, and sample characteristics. Finally, they also underscore the importance of examining more fine-grained personality constructs to better understand individual variability in the extent to which foreign language use shapes moral judgment.

Highlights

- Personality traits modulate the foreign language effect (FLe) on moral decisions
- Lower conscientiousness and higher emotional stability reduce FLe susceptibility
- Exception: FLe overrides deontological tendencies linked to lower extraversion
- Findings support fine-grained trait analysis over broad personality measures in FLe

1. Introduction

Recent research highlights how non-native language use influences decision-making and reasoning. This ‘foreign language effect’ (FLe) notably affects moral decision-making, with foreign language use increasing utilitarian choices in dilemmas balancing common good versus moral rules (see Costa et al. (2014) for foundational work). While a few studies have reported null effects (e.g., Mills & Nicoladis, 2023; Muda et al., 2020; Yavuz et al., 2024), meta-analytic evidence indicates that the FLe is reliable across moral contexts (Circi et al., 2021; Costa et al., 2019; Del Maschio et al., 2022). An example of the type of dilemma where the FLe has been typically observed is the footbridge variant (Thomson, 1976) of the trolley problem (Foot, 1967), where a runaway trolley threatens five workers. Diverting the trolley’s path by pushing a large man onto the tracks can save the five workers at the expense of the man’s life. This decision reflects utilitarian thinking (greater in individuals responding in their foreign language), while refraining from such action aligns with deontological principles (Conway et al., 2018).

The FLe has been associated with the idea that non-native languages often lack the emotional grounding typically developed through early-life experiences. Several mechanisms have been proposed to explain why moral judgments may differ when using a foreign language. One account emphasizes increased analytical processing in FL contexts, which may promote more utilitarian choices by downregulating intuitive, emotion-based responses (Costa et al., 2014). Another explanation suggests that FL use reduces emotional resonance, thereby weakening the instinctive aversion to causing harm (Geipel et al., 2015a; Hayakawa et al., 2017). Additionally, reduced access to internalized social and moral norms has been proposed as a factor that could shift decisions in FL contexts (Geipel et al., 2015b). Finally, FL contexts may foster greater psychological distance from the moral dilemma, encouraging a more detached and deliberative evaluation (Braida et al., 2023; Corey et al., 2017; Ivaz et al., 2016; Shin & Kim, 2017). Although these explanations focus on different mechanisms, they generally agree that FL contexts attenuate affective and social intuitions that typically drive deontological responses – a view supported

by recent evidence favoring emotional blunting over increased deliberation (Bialek et al., 2019; Hayakawa et al., 2017). However, while significant, the effect size has been found to be small-to-moderate (e.g., Circi et al., 2021; Del Maschio et al., 2022). This prompts the question: Are there individual differences that moderate the FLe? In this article, we explore the potential moderating role of basic personality factors on the FLe. We hypothesize that one reason the FLe is generally modest is that the effects of an FL cannot override particularly strong deontological positions influenced by certain personality factors. To test this hypothesis, we focus on the Big Five personality factors, a widely used framework for capturing individual differences. To elicit moral decisions, we employ dilemmas modeled after the classic trolley problem, which are specifically designed to contrast utilitarian and deontological responses (a contrast often referred to as the ‘traditional’ moral dilemma juxtaposition).

1.1. *The Big Five factors and their association with moral judgment*

The Big Five model offers a widely accepted framework for organizing personality into five broad domains – extraversion, agreeableness, conscientiousness, neuroticism, and openness – enabling researchers to identify meaningful patterns rather than isolated traits and providing a shared vocabulary for describing personality (John et al., 2008). Although these dimensions do not capture every nuance of individuality, each reflects a constellation of related characteristics that summarize major aspects of personality variation. For instance, extraversion reflects an energetic, outward-oriented style marked by sociability, assertiveness, activity, and positive emotionality (with low scores indicating introversion; John & Srivastava, 1999). Agreeableness captures a prosocial orientation toward others, including empathy, trust, modesty, and a desire for harmony. Conscientiousness reflects self-regulation in pursuit of goals, encompassing planning, organization, perseverance, norm-following, and delayed gratification. Neuroticism, by contrast, refers to emotional instability and the tendency to experience negative affect such as anxiety, sadness, or irritability. Finally, openness describes cognitive and experiential breadth, including curiosity, imagination, aesthetic sensitivity, and a preference for novelty and complexity.

Together, these five dimensions offer a comprehensive framework for distinguishing individual differences in behavior. Their relevance has been supported by consistent associations between Big Five inventory scores and patterns of self-reported everyday actions (John et al., 2008). For example, individuals high in extraversion are more likely to introduce themselves to strangers at social events or take initiative in group projects, whereas those lower in this trait (commonly referred to as introverts) tend to avoid voicing disagreement and may experience more interpersonal difficulties. People high in agreeableness often speak positively about others, offer help or lend personal items, and comfort friends in distress. Conscientious individuals typically arrive punctually, work diligently to achieve top academic results, and check their written work carefully, while those scoring lower may leave tasks like washing dishes unattended. Those low in neuroticism – that is, more emotionally stable – tend to accept situations without complaint and remain relaxed under stress, whereas those higher on this trait may feel easily hurt or unsettled when faced with conflict. Finally, individuals high in openness often pursue learning for its own sake, watch educational content, rearrange their environment creatively, or seek out novel and stimulating activities to break routine.

Unlike the well-established links between personality and everyday behavior illustrated above, evidence connecting the Big Five to

moral inclinations is limited, hindering predictions about their role in moderating the FLe. Most studies rely on self-report measures rather than behavioral tasks. For instance, Abbasi-Asl and Hashemi (2019) found that agreeableness predicted moral sensitivity and identity, conscientiousness predicted moral identity and courage, while neuroticism and extraversion negatively predicted some moral components. Similarly, Rengifo and Laham (2022) showed that lower agreeableness and conscientiousness predicted greater moral disengagement, linked to justifying unethical behavior, whereas Khan et al. (2016) found that conscientiousness and agreeableness were positively related to idealism, and neuroticism to relativism.

More closely aligned with our approach, Luke and Gawronski (2022a, 2022b) used moral dilemmas instead of self-report measures. Nevertheless, their work was conducted entirely in a single-language context and did not examine the role of foreign language use in moral decision-making, and their primary aim was to shed light on how personality factors might modulate moral choices. Despite this difference in scope with our study, their findings provide an important point of reference, as they represent one of the few studies to examine how basic personality traits relate to moral choices elicited by moral dilemmas rather than self-report measures. They found that while agreeableness and openness increased consequence sensitivity, extraversion decreased it. Moral norm sensitivity correlated positively with agreeableness, conscientiousness, and openness, and negatively with neuroticism. Preference for inaction was predicted positively by extraversion, agreeableness, conscientiousness, and openness, and negatively by neuroticism. However, these results replicated only partially across different samples, highlighting the limitations of relying solely on undergraduate participants.

Notably, in this limited body of prior research, proposed links between specific personality traits and components of moral reasoning remain theoretically open and are often compatible with multiple interpretations, reflecting the complexity of mapping broad personality dimensions onto moral decision processes. Moreover, most of these studies have not employed the classical deontological–utilitarian contrast that we adopt here, given its widespread use in the FLe literature. This further complicates the task of generating precise predictions about which traits might support deontological inclinations. In this regard, the only direct reference, to our knowledge, is the ‘traditional dilemma score’ computed by Luke and Gawronski (2022a, 2022b) alongside the three parameters of the CNI model (sensitivity to consequences, moral norms, and action preferences). This additional score was derived from a subset of moral dilemmas similar in structure to the classic trolley problem – featuring actions with a clear net benefit and aligned with moral norms – rather than the broader set of dilemmas used to estimate sensitivity to consequences, norms, and action bias in the CNI framework. Luke and Gawronski’s (2022b) findings based on the traditional dilemma score revealed some inconsistencies across studies and analytical strategies (zero-order correlations vs. multiple regressions) as well as test–retest phases. Nevertheless, correlational analyses consistently showed that agreeableness, conscientiousness, and openness were associated with more deontological responses. Yet, when controlling for the shared variance among traits in multiple regression analyses, only conscientiousness emerged as a reliable predictor of deontological inclinations. This latter result – reported in Luke and Gawronski (2022b) – is arguably more robust, as regression models account for confounding effects among predictors, much like the mixed-effects models used in our study.

In summary, the present study is theoretically motivated by the observation that the FLe is not uniform across individuals,

suggesting that stable dispositional factors, such as personality traits, may influence its expression. Specifically, we hypothesize that traits favoring deontological tendencies could modulate the FLe. However, the precise ways in which personality traits affect moral decisions remain largely unknown, as prior studies employing diverse methodologies and measures have produced overall inconsistent results. The only prior work using moral dilemmas, Luke and Gawronski (2022a, 2022b), identified conscientiousness as the most robust predictor of deontological responding once shared variance among traits was controlled. In the present study, we aim to identify which personality traits are associated with deontological responding under our methodological conditions and, simultaneously, test our main goal: whether these traits modulate the FLe across two experimental language groups. Guided by Luke et al., we tentatively expect conscientiousness to relate to deontological responding – and hence modulate the FLe – while also exploring all other traits without strong directional predictions, remaining open to the possibility that additional or different traits could play a role.

1.2. Typical FLe pattern in ‘traditional’ moral dilemmas: A methodological baseline

To replicate the typical FLe in moral decision-making and to test whether personality traits modulate this effect, we used traditional sacrificial dilemmas such as the trolley problem. These dilemmas instantiate a clear utilitarian–deontological contrast and have been used extensively in prior FLe research, where the effect has been most consistently observed (e.g., Corey et al., 2017; Costa et al., 2014). Although traditional dilemmas have been criticized for their limited ecological validity and for omitting additional dimensions formalized in models such as the CNI framework (Gawronski et al., 2017), they reliably elicit strong deontological intuitions and robust individual variability in moral responding. Their use here therefore reflects a deliberate methodological choice to maximize the likelihood of replicating the FLe observed in prior studies, which is essential for assessing whether personality traits modulate this effect.

Traditional moral dilemmas are often categorized as personal or impersonal, which elicit different psychological and emotional responses. Personal dilemmas, like the footbridge trolley problem, require directly harming an individual, while impersonal dilemmas, such as the switch version, involve indirect harm by redirecting threats. Typically, about 80% of responses to impersonal dilemmas are utilitarian, compared to less than 50% for personal ones (Carmona-Perera et al., 2015). This difference reflects the stronger emotional engagement in personal dilemmas due to direct involvement in causing harm (Greene et al., 2004, 2008).

Personal moral dilemmas can vary in conflict intensity, typically classified as high or low conflict. High-conflict dilemmas – such as the ‘vaccine’ scenario, where participants must decide whether to sacrifice one person to save millions with a vaccine – elicit strong tension between utilitarian outcomes and individual harm. These dilemmas are associated with approximately 50% utilitarian responses, longer decision times, and low consensus (Carmona-Perera et al., 2015; Greene et al., 2004). In contrast, low-conflict dilemmas – such as the ‘transplant’ case, where participants consider taking organs from a healthy individual (resulting in their death) to save five others – typically yield fewer than 10% utilitarian responses and faster decisions, reflecting a more clear-cut aversion to harm (Carmona-Perera et al., 2015).

Of note, the FLe has been predominantly observed in personal dilemmas, whether high or low conflict, while it appears to be absent or less pronounced in impersonal dilemmas. For example, different studies have shown that the FLe is typically observed in the footbridge version of the trolley dilemma but not in the switch version (e.g., Cipolletti et al., 2016; Costa et al., 2014; Geipel et al., 2015a). Similarly, in an extensive series of experiments, Corey et al. (2017) found that in only 1 out of 10 experiments using an impersonal switch-like dilemma was there evidence of the FLe, whereas in 7 out of 10 personal footbridge-like dilemmas, the FLe was evident. Replicating the well-established response patterns observed in traditional moral dilemmas, as well as in the FLe literature, will provide a solid empirical context from which to examine our central question: whether individuals with personality traits favoring deontological tendencies maintain these leanings when moral decisions are made in a foreign language.

2. Methods

To investigate whether Big Five personality traits modulate the FLe in moral decision-making, participants completed a set of traditional moral dilemmas (BrMoD; Carmona-Perera et al., 2015) in either their native (Spanish) or a foreign language (English), along with the Mini-IPIP-PW (Martínez-Molina & Arias, 2018) to assess personality in their native language. To determine whether potential confounds needed to be controlled for, we also measured reasoning style using the Cognitive Reflection Test (CRT; Sirota & Juanchich, 2018) and general intelligence with Raven’s Advanced Progressive Matrices (RAPM; Raven et al., 1998).

2.1. Participants

The sample consisted of 236 female participants, who gave written consent to participate in exchange for monetary compensation (approximately £6.00). The study was granted ethical approval by the Bioethical Committee of the University of Barcelona (Institutional Review Board 00003099). However, given evidence that women and men may differ in personality traits relevant to moral judgment – particularly emotional stability, which tends to be lower in women across several large-scale studies (Costa et al., 2001; Weisberg et al., 2011) – and considering funding constraints that limited our ability to recruit a sufficiently powered, gender-balanced sample, we chose to focus on one gender. Since many prior FLe studies have predominantly included female participants – including foundational work (e.g., Cipolletti et al., 2016; Corey et al., 2017; Geipel et al., 2015a [studies 1 & 3], 2015b; Geipel et al., 2016; Hadjichristidis et al., 2015) – we opted to include only women, ensuring comparability and consistency with the existing literature. Additionally, women may experience greater psychological connectedness to emotionally charged dilemmas, potentially reducing the psychological distance typically induced by foreign language use (Ivaz et al., 2016; Shin & Kim, 2017). While this was a secondary consideration, it further supported our decision to focus exclusively on female participants.

All participants were native Spanish speakers from Spain. Importantly for the purposes of the study, and to ensure greater homogeneity in cultural background and linguistic exposure, individuals of Latin American origin were not included. Participants were randomly assigned to one of two groups: 118 participants completed the study in Spanish (NL group), while another 118 completed the study in English (FL group). There were slight between-group differences in age (NL group: $Mdn = 25$, $IQR = 16.75$,

range = 18–57; FL group: $Mdn = 23$, $IQR = 7$, range = 18–58; $W = 8091$, $p = 0.030$; Cliff's delta = 0.16, 95% CI [0.015, 0.303], small effect). To improve generalizability beyond student samples (Luke & Gawronski, 2022a), we included participants from a broader range of educational attainment. The majority of participants had completed their studies up to the bachelor's degree level, with 61.02% in the NL group and 55.93% in the FL group. Nevertheless, a small but significant difference was found in educational attainment between the groups ($W = 4890$, $p = 0.003$; Cliff's delta = -0.21 , 95% CI $[-0.334, -0.070]$, small effect). This difference was attributed to a higher prevalence of education levels beyond a bachelor's degree (master's or doctoral) in the FL group (33.9%) compared to the NL group (19.49%), whereas education levels below a bachelor's degree (secondary school, vocational training) were more common in the NL group (19.49%) than in the FL group (10.17%).

Some participants in both groups were native to Spanish autonomous communities with co-official languages – primarily Catalonia, and to a lesser extent the Basque Country, Valencian Community, and Galicia. There were no significant between-group differences in the number of participants with a co-official regional L2 ($p = 0.154$) (Table 1A). Both groups also had comparable English language backgrounds, with no significant between-group differences in the context of first exposure (all $p_s > 0.067$) (Table 1B). Most participants reported an age of acquisition (AoA) around 8 years, with no significant differences between groups. Although a few were exposed earlier, they lacked regular home input and did not attain native-like proficiency, as reflected in their self-rated scores. English proficiency was assessed in both groups via a self-report questionnaire covering speaking, listening, reading, and writing (Likert scale: 1 = low, 7 = high); no significant group differences were found across individual skills or in the composite score (all $p_s > 0.12$). LexTALE scores, originally available only for the FL group, indicated proficiency levels corresponding to B2–C2 on the Common European Framework (CEF) (Table 1C; Lemhöfer & Broersma, 2012). To provide an additional objective check of English proficiency comparability between groups, NL participants were subsequently recontacted to complete the LexTALE, yielding data from 48 participants. Their scores ($M = 83.3$, $SD = 9.76$) were descriptively similar to those of the FL group ($M = 79.33$, $SD = 10.83$). A Bayesian independent-samples t-test using a default Jeffreys–Zellner–Siow (JZS) prior yielded anecdotal evidence for a group difference ($BF_{10} = 1.70$), indicating that the data do not provide strong support for either the presence or absence of a difference in English lexical proficiency between groups. The reduced NL sample reflects the post hoc nature of this assessment and should therefore be interpreted as a robustness check rather than a core group comparison. Together with self-reported proficiency and age-of-acquisition measures, these results support the characterization of both groups as bilingual English speakers, making it unlikely that any observed differences in moral decision-making can be attributed to monolingualism versus bilingualism.

2.2. Materials and procedure

All participants completed the BrMoD battery of traditional moral dilemmas (Carmona-Perera et al., 2015). Prior to this, they responded to a series of non-moral dilemmas, also drawn from Carmona-Perera et al. (2015), which served to screen for inattentive or random responding. Following the moral dilemmas, participants completed the CRT (reasoning style), the Mini-IPIP-PW (Big Five personality traits), and the RAPM (general intelligence), in that order. The entire procedure was implemented using Qualtrics and administered via Prolific.

Table 1. Language background of participants

A. Number of NL and FL participants with a regional co-official L2			
	NL (N: 118)	FL (N: 118)	Chi-square test
Total	38	50	$\chi^2(1) = 2.03$, $p = 0.154$
Catalan	23	36	
Galician	4	7	
Basque	6	5	
Aragonese	1	0	
Asturian	1	0	
Valencian	3	2	
B. Context of first English exposure			
	NL [(N: 118)]	FL [N: 118]	Chi-square test
Home	12	23	$\chi^2(1) = 3.35$, $p = 0.067$
School	102	110	$\chi^2(1) = 2.27$, $p = 0.132$
Other	34	37	$\chi^2(1) = 0.08$, $p = 0.777$
C. English proficiency and age of acquisition (AoA)			
	NL [$\bar{x}(SD)$]	FL [$\bar{x}(SD)$]	Wilcoxon test
Self-rated English proficiency levels in NL and FL participants			
Reading	5.91 (1.25)	6.11 (0.86)	$W = 6611$, $p = 0.477$
Writing	5.31 (1.43)	5.51 (1.02)	$W = 6607$, $p = 0.483$
Speaking	5.04 (1.68)	5.43 (1.33)	$W = 6170$, $p = 0.122$
Listening	5.92 (1.28)	6.08 (1.07)	$W = 6606$, $p = 0.471$
Composite score	5.24 (1.31)	5.45 (0.92)	$W = 6457$, $p = 0.317$
LexTALE score	–	79.33 (10.83)	–
English AoA	7.5 (6.15) range: 0–30; IQR: 8.0	7.9 (5.23) range: 0–28; IQR: 7.5	$W = 6358$, $p = 0.249$

N = adds up to more than the total sample because several participants indicated exposure to more than one place where they learned English. Other = audio-visual media, social media, job, academies, etc. LexTALE score = ((number of words correct/40*100) + (number of nonwords correct/20*100)) / 2. Composite score = the average across the four proficiency domains. IQR = Interquartile range.

Non-moral dilemmas (Carmona-Perera et al., 2015): These logic-based dilemmas assess rational decision-making without moral content. Because responses are typically straightforward, low scores may indicate random or inattentive responding (Lange et al., 2010). Following Carmona-Perera et al., participants were excluded if they answered fewer than six dilemmas correctly (i.e., <94% accuracy). Eleven participants (3 NL, 8 FL) failed to meet this criterion and were excluded from the final sample ($N = 236$). An example is the ‘broken videocamera’ dilemma, where one must choose between repairing an old device or buying a better one for the same price.

Traditional moral dilemmas (BrMoD battery; Carmona-Perera et al., 2015): There were 21 hypothetical scenarios that require a ‘yes/no’ decision about causing harm to someone for the sake of a

Table 2. Description of the traditional moral dilemmas (BrMoD battery, Carmona-Perera et al., 2015)

Impersonal dilemmas
Standard fumes. A hospital watchman must decide whether to redirect deadly fumes from a room containing three patients to one with a single patient.
Vaccine policy. A government health official must decide whether to endorse the use of a newly developed, highly effective vaccine, despite the risk that a few individuals may contract the disease as a result of vaccination.
Environmental policy A1. A government legislator must choose between Policy A, which carries a 10% chance of causing 1,000 deaths, and Policy B, which entails an 8% risk of causing 10,000 deaths.
Environmental policy A2. A government legislator must choose between Policy A, which carries a 10% chance of causing 1,000 deaths, and Policy B, which entails a 12% risk of causing 10 deaths.
Guarded speed boat. An individual faces the decision of whether to deceive a vigilant guard to commandeer the boat he is guarding, aiming to reach stranded tourists facing an imminent storm.
Trolley – Switch version. A person must decide whether to divert a runaway trolley towards a single railway worker to prevent the deaths of five others.
Personal dilemmas
High conflict
Trolley – Footbridge version. A person on a footbridge must decide whether to push a stranger onto the tracks to stop a runaway trolley and save the lives of five workmen.
Lifeboat modified. On a sinking boat, a person faces the dilemma of sacrificing one injured person's life to save the rest of the passengers.
Ecologists. A hostage within an ecological group held captive by terrorists must contemplate sacrificing one fellow hostage to secure the escape of the remaining captives, which includes eight children and herself.
Lawrence of Arabia. As a leader of an army unrelated to any tribe but overseeing members of two distinct tribes, one must decide whether to execute a member of one tribe who has slain a member of the other tribe, thereby averting a catastrophic war between the two factions.
Sophie's choices. To prevent both of her children from enduring painful experiments by the enemy, a mother faces the dilemma of sacrificing one child to save the other.
Sacrifice. After inadvertently trespassing on sacred ground, a mother must decide whether to sacrifice her oldest son to save the rest of her family, as demanded by the local clan's leader.
Vitamins. Faced with the need to save the lives of six individuals with a genetic vitamin deficiency, the leader of a stranded mountaineering expedition must decide whether to forcibly extract a kidney from a member of the group, knowing that although he will not die, his health will be compromised.
Vaccine test. Faced with a global epidemic, a researcher must decide whether to sacrifice one person's life by administering a potentially deadly injection to identify a life-saving substance.
Euthanasia. An injured soldier faces imminent capture and torture. He pleads not to be left behind, but assisting him will result in the capture of the entire group. The leader must decide whether to prevent the soldier's torture by shooting him herself.
Low conflict
Transplant. A doctor must decide whether to transplant organs from an unwilling healthy donor, leading to his death but ensuring the survival of five dying patients.
Country road. A driver must decide whether to assist a hiker with injured legs knowing that it would likely result in blood staining the car's leather upholstery.
Architect. An architect must decide whether to push their despised boss off a building to rid themselves of his influence, knowing it would be perceived as an accident.
Smother for dollars. A person must decide whether to smother a terminally ill man with a pillow in return for a large sum of money from the man's son who needs his father to die before the substantial life insurance expires.
Infanticide. A teenage girl, pregnant and unprepared to care for a newborn, faces the decision of whether to dispose of the baby in a dumpster behind the school.
Plane crash. Stranded in the Himalayas with limited resources and a young boy with a broken leg, a survivor must decide whether to sacrifice the boy to increase her own and another adult survivor's chances of survival.

greater good. These scenarios were presented in first person, prompting participants to imagine themselves as the main character of the story. The scenarios consisted of 6 impersonal dilemmas and 15 personal dilemmas, of which 6 were low conflict and 9 were high conflict (see Table 2 for a description of the dilemmas included in each category)¹. We followed the methodology established by

Carmona-Perera et al. (2015), using a random order of dilemmas administered to all participants. Participants were instructed that each question represented a personal decision and had no right or wrong answer. At the end of each dilemma, participants were prompted with a question about the hypothetical action, expressed in the following format: 'Would you...in order to ...?' Participants responded by selecting either the 'yes' or 'no' button to indicate their answer, with 'yes' responses always indicating the utilitarian choice (i.e., the commission of the proposed action). There was no time limit for reading the scenario, and upon pressing the button, the next dilemma appeared on the screen automatically.

¹The BrMoD battery contains one further high-conflict personal dilemma known as the 'Crying baby'. However, due to a technical issue, this dilemma was not administered to any participant.

Cognitive reflection test (CRT, multiple-choice version, Sirota & Juanchich, 2018): This seven-item multiple-choice test assessed participants' ability to override intuitive but incorrect responses in favor of reflective reasoning. Each item had four options: the correct answer (1 point; *CRT-reflective score*), the intuitive error (1 point; *CRT-intuitive score*), and two distractors. To avoid FLE-related bias, all participants completed the CRT in their native language. There were no significant differences between the NL and FL groups in either the CRT-reflective score (NL group's *Mdn* = 3 (range = 0–7), FL group's *Mdn* = 2.5 (range = 0–7); $W = 6846.5$, $p = 0.824$) or the CRT-intuitive score (NL group's *Mdn* = 3 (range = 0–7), FL group's *Mdn* = 3 (range = 0–6); $W = 7549$, $p = 0.256$, 95%). In other words, the NL and FL groups demonstrated equivalent reasoning styles.

Mini international personality item pool-five-factor model-positively worded (Mini-IPIP-PW; Martínez-Molina & Arias, 2018): Personality was assessed using the Mini-IPIP-PW (Martínez-Molina & Arias, 2018), a positively worded 20-item questionnaire measuring the Big Five traits: extraversion, agreeableness, conscientiousness, emotional stability (i.e., the inverse of neuroticism, with lower values indicating higher neuroticism), and openness. Each trait was measured with four items rated on a 5-point Likert scale (1 = totally disagree, 5 = totally agree), yielding scores from 4 to 20 per trait. The questionnaire was administered in participants' native language to minimize psychological distancing effects associated with foreign language use (Ivaz et al., 2016). In our sample, internal consistency was satisfactory to good across all five traits – extraversion ($\alpha = .79$), agreeableness ($\alpha = .85$), conscientiousness ($\alpha = .77$), emotional stability ($\alpha = .82$), and openness ($\alpha = .77$) – with an overall $\alpha = .75$. Participants in the FL group displayed slightly higher levels of extraversion compared to those in the NL group (NL group's *Mdn* = 9 (range = 4–17), FL group's *Mdn* = 10 (range = 4–20); $W = 5904$, $p = 0.043$). No significant differences were observed between the two groups in the other personality factors, including agreeableness (NL group's *Mdn* = 16 (range = 4–20), FL group's *Mdn* = 16 (range = 4–20); $W = 7029$, $p = 0.898$), conscientiousness (NL group's *Mdn* = 13.5 (range = 5–20), FL group's *Mdn* = 13 (range = 5–20); $W = 6891$, $p = 0.893$), emotional stability (NL group's *Mdn* = 9 (range = 4–19), FL group's *Mdn* = 9 (range = 4–20); $W = 7060.5$, $p = 0.851$), and openness (NL group's *Mdn* = 14 (range = 6–20), FL group's *Mdn* = 13 (range = 6–20); $W = 7615.5$, $p = 0.211$).

Ravens advanced progressive matrices (Superior Scale I; Raven et al., 1998): This 12-item task assessed general intelligence by requiring participants to identify the missing piece in visual patterns from eight options. To avoid FLE-related bias, instructions were provided in participants' native language (NL), regardless of group. Participants had 10 minutes to complete the task, and total correct responses constituted the Raven score. There was no significant difference in the Raven score between the two groups (NL group's *Mdn* = 10, (range = 4–12) FL group's *Mdn* = 10 (range = 4–12); $W = 6118$, $p = 0.101$).

3. Data analyses

We employed generalized linear mixed models using likelihood ratio tests to compare the relative fit of two nested models: a full model and a reduced model excluding a single predictor. A significant difference in model fit indicated that the omitted predictor contributed uniquely to explaining variance in the dependent variable. We adopted a backward elimination procedure, beginning with a comprehensive model that included all fixed effects and

iteratively comparing it to reduced models, each omitting one predictor. Interactions were assessed using the same approach – that is, by comparing a model with the interaction term to an otherwise identical model without it. In line with common statistical practice, predictor inclusion in the final model was based on the results of these comparisons (i.e., ANOVA likelihood ratio tests). In addition to likelihood ratio tests, we report the Akaike Information Criterion (AIC) for each model as an index of model fit, with lower AIC values indicating better fit. Differences in AIC (Δ AIC) can be interpreted in terms of practical relevance: a Δ AIC >2 suggests meaningful improvement, with values around 2–4 indicating small, 4–7 moderate, and > 10 substantial improvement in model fit (Burnham & Anderson, 2002). Parameter estimates and associated statistics for the individual predictors are not reported for intermediate models but are presented exclusively for the final models retained after model comparison. These details, including standardized effect sizes (β -values), reflect both the direction and magnitude of these effects and are shown in Table 3.

Table 3. Characteristics of the final models

A. Results for all data (including both personal and impersonal dilemmas)					
	Estimate	SE	t-value	p-value	β -value
(Intercept)	2.223	0.688	3.23	0.001	1.93
Age	−0.017	0.008	−2.115	0.034	−0.12
Language group					
NL vs. FL	0.094	0.193	0.487	0.626	0.03
Type of dilemma					
Personal vs. impersonal	−3.711	0.816	−4.55	0.0001	−2.75
Extraversion	0.046	0.022	2.088	0.037	0.06
Emotional stability	−0.043	0.024	−1.825	0.068	−0.07
Conscientiousness	0.06	0.022	2.709	0.007	0.11
Interaction					
Language group*type of dilemma	0.515	0.233	2.209	0.027	0.31
B. Results for personal moral dilemmas					
	Estimate	SE	t-value	p-value	β -value
(Intercept)	−0.129	0.371	−0.348	0.728	−0.15
Age	−0.019	0.01	−1.939	0.050	−0.16
Group					
NL vs. FL	0.572	0.172	3.327	< 0.001	0.44
Level of conflict					
Low vs. high	−3.508	0.588	−5.967	< 0.001	−2.52
Extraversion	0.125	0.034	3.395	< 0.001	0.21
Emotional stability	−0.071	0.028	−2.551	0.012	−0.09
Conscientiousness	0.07	0.028	2.545	0.011	0.13
Interaction					
Group*extraversion	−0.101	0.05	−2.012	0.044	−0.16

Note: A. Results of the final model including the interaction between 'language group' and 'type of dilemma' (all data). B. Results of the final model performed on the data of personal moral dilemmas. β -values: standardized effect sizes for fixed effects (≈ 0.1 : small, ≈ 0.3 : moderate, > 0.5 : large; Cohen, 1988).

However, as interpreting individual fixed-effect parameters in linear mixed models can be misleading in isolation (e.g., Luke, 2017), their contribution should be considered in the context of overall model fit.

We applied this procedure in three separate analyses. The initial model (Section 4.1) included both personal and impersonal dilemmas. Given previous findings suggesting that the FLE tends to be more salient in personal dilemmas (e.g., Cipolletti et al., 2016; Corey et al., 2017; Costa et al., 2014; Geipel et al., 2015a), we explicitly tested for an interaction between ‘type of dilemma’ (personal vs. impersonal) and ‘language group’ (native vs. foreign). For the sake of clarity, we note here that this interaction was significant and motivated the two subsequent analyses: one focusing exclusively on personal dilemmas (Section 4.2) and the other on impersonal dilemmas (Section 4.3).

Across all three analyses, the binary dependent variable was ‘type of decision’ (utilitarian vs. deontological). This variable encoded the response provided by each participant for every moral dilemma. For instance, opting to push the large man onto the tracks in the footbridge version of the trolley dilemma represents a utilitarian decision, while responding negatively to the same dilemma reflects a deontological perspective. In the initial analysis (Section 4.1), this variable collapsed responses across both personal and impersonal dilemmas, whereas in the subsequent analyses, it was based solely on responses to either personal (Section 4.2) or impersonal (Section 4.3) dilemmas.

All three analyses included both a fixed-effect structure and a random-effect structure. The fixed-effect structure comprised **experimental predictors** – those central to our research questions – and **control predictors**, included to account for potential confounding factors.

The experimental predictors were ‘type of dilemma’ (personal vs. impersonal), ‘language group’ (native vs. foreign), and all Big Five factors of personality (‘extraversion’, ‘agreeableness’, ‘conscientiousness’, ‘emotional stability’, and ‘openness’). Note that ‘type of dilemma’ was not included in the two separate analyses by dilemma type (i.e., the personal dilemma analysis in Section 4.2 and the impersonal dilemma analysis in Section 4.3), and that the analysis of personal moral dilemmas also included ‘level of conflict’ (high vs. low). Aside from these differences, the set of experimental predictors was consistent across all three analyses.

To minimize multicollinearity, we limited the number of control predictors to those variables that (a) significantly differed between the NL and FL groups (Miller & Chapman, 2001; Pedhazur, 1997; Tabachnick & Fidell, 2019) or (b) showed significant correlations with the percentage of utilitarian responses within each group, based on Spearman correlations (and a point-biserial correlation for the binary bilingual status variable) (Field et al., 2012; Tabachnick & Fidell, 2019; West et al., 2015). The variables considered in these correlations included all sociolinguistic measures (age, educational attainment, English AoA, self-rated speaking, listening, reading, and writing skills, as well as the composite score from the self-reported English proficiency questionnaire), the LexTALE score (FL group only), bilingual status (i.e., whether the participant was a speaker of a co-official regional language), and the reasoning and general intelligence measures (CRT-reflective score, CRT-intuitive score, and Raven score). Based on these criteria, only ‘age’ and ‘educational attainment’ were included as control predictors in the fixed-effect structure, as these variables showed significant between-group differences. In addition, ‘age’ was the only variable that significantly correlated with the percentage of utilitarian responses ($\rho = -0.25$, $p = 0.015$; no other variable

showed significant correlations with moral decision-making; all $\rho_s > 0.11$, all $p_s > 0.196$).

Following the recommendations of Barr et al. (2013), we employed the maximum random-effect structure justified by the design. This entailed the inclusion of ‘participant’ and ‘dilemma’ as random effects, along with by-participant random slopes for only the experimental predictors.

The analyses were conducted with the R package (version 4.0.2; R Development Core Team, 2020).

4. Results

4.1. Initial analysis

This initial analysis included the experimental predictors ‘type of dilemma’ (personal vs. impersonal), ‘language group’ (native vs. foreign), and all Big Five factors, as well as the control predictors ‘age’ and ‘educational attainment’. The binary dependent variable, ‘type of decision’ (utilitarian vs. deontological), collapsed responses across both personal and impersonal dilemmas.

The full model (AIC = 4199.7) provided a significantly better fit than the reduced models when testing the effects of ‘language group’ ($\chi^2(1) = 10.13$, $p = 0.001$; AIC reduced model = 4207.9, $\Delta\text{AIC} = 8.2$), ‘type of dilemma’ ($\chi^2(1) = 13.58$, $p < 0.001$; AIC reduced model = 4211.3, $\Delta\text{AIC} = 11.6$), ‘age’ ($\chi^2(1) = 4.67$, $p = 0.031$; AIC reduced model = 4202.4, $\Delta\text{AIC} = 2.7$), ‘extraversion’ ($\chi^2(1) = 4.599$, $p = 0.032$; AIC reduced model = 4202.3, $\Delta\text{AIC} = 2.6$), ‘emotional stability’ ($\chi^2(1) = 4.835$, $p = 0.028$; AIC reduced model = 4202.6, $\Delta\text{AIC} = 2.9$), and ‘conscientiousness’ ($\chi^2(1) = 7.384$, $p = 0.007$; AIC reduced model = 4205.1, $\Delta\text{AIC} = 5.4$). These main effects revealed that utilitarian decisions decreased with age and higher levels of emotional stability, while they were more frequent when using an FL, as well as with higher levels of extraversion and conscientiousness. No significant difference between the full model and the reduced models was observed when testing the effects of ‘educational attainment’ ($\chi^2(1) < 1$), ‘agreeableness’ ($\chi^2(1) < 1$), or ‘openness’ ($\chi^2(1) = 1.189$, $p = 0.276$).

Next, we examined the interaction between ‘type of dilemma’ and ‘language group’. The full model incorporating this two-way interaction provided a significantly better fit than the model without it ($\chi^2(1) = 4.09$, $p = 0.043$; AIC full model with interaction = 4197.6, $\Delta\text{AIC} = 2.1$), indicating that participants in the FL group showed a greater contrast in utilitarian judgments between personal and impersonal dilemmas than those in the NL group (see Table 3A for the direction of effects based on factor level coding). Consequently, we conducted subsequent analyses separately for personal (Section 4.2) and impersonal (Section 4.3) dilemmas.

Table 3A presents the characteristics of the final model containing all the significant predictors and the two-way interaction between ‘language group’ and ‘type of dilemma’, along with β -values (standardized effect sizes).

4.2. Analysis of personal moral dilemmas

This analysis included the experimental predictors ‘level of conflict’ (high vs. low), ‘language group’ (native vs. foreign), and all Big Five factors, as well as the control predictors ‘age’ and ‘educational attainment’. The binary dependent variable, ‘type of decision’ (utilitarian vs. deontological), included only responses to personal dilemmas.

The full model (AIC = 3118.8) provided a significantly better fit than the reduced models when testing the effects of ‘language

group' ($\chi^2(1) = 15.05, p < 0.001$; AIC reduced model = 3131.9, $\Delta AIC = 13.1$), 'age' ($\chi^2(1) = 4.91, p = 0.027$; AIC reduced model = 3121.7, $\Delta AIC = 2.9$), 'extraversion' ($\chi^2(1) = 8.47, p = 0.004$; AIC reduced model = 3125.3, $\Delta AIC = 6.5$), 'emotional stability' ($\chi^2(1) = 9.71, p = 0.002$; AIC reduced model = 3126.5, $\Delta AIC = 7.7$), and 'conscientiousness' ($\chi^2(1) = 7.54, p = 0.006$; AIC reduced model = 3124.4, $\Delta AIC = 5.6$), as well as 'level of conflict' (high vs. low) ($\chi^2(1) = 19.07, p < 0.001$; AIC reduced model = 3135.9, $\Delta AIC = 17.1$). No significant difference between the full model and the reduced models was observed when testing the effects of 'educational attainment' ($\chi^2(1) < 1$), 'agreeableness' ($\chi^2(1) < 1$), or 'openness' ($\chi^2(1) = 3.56, p = 0.060$).

These main effects indicated that utilitarian decisions in personal dilemmas decreased with age and higher levels of emotional stability. In contrast, utilitarian decisions increased when using an FL. Consistent with most prior studies, this increase remained in a modest (albeit significant) range, as the average utilitarianism – computed from the individual percentages of utilitarian decisions ('yes' responses) – was 8.76% higher when using an FL compared to an NL. Utilitarian decisions also increased with higher levels of extraversion and conscientiousness. Furthermore, the results revealed that, compared to low-conflict personal dilemmas, high-conflict personal dilemmas led to an increase in utilitarian decisions, with an imbalance in average utilitarianism between low-conflict (8.47%) and high-conflict (52.02%) personal dilemmas replicating findings from prior studies (e.g., Carmona-Perera et al., 2015).

Additionally, we examined whether 'language group' interacted with any experimental or control predictor. The full model incorporating the two-way interaction between 'language group' and 'extraversion' provided a significantly better fit than the model without it ($\chi^2(1) = 4.46, p = 0.035$; AIC full model with interaction = 3116.4, $\Delta AIC = 2.4$). No other full model incorporating the interaction between group and another predictor differed from the full model with no interaction (all $p_s > 0.198$), indicating that 'language group' interacted with no predictor except for 'extraversion'. To interpret the significant interaction between 'extraversion' and 'language group', we conducted a follow-up analysis estimating the simple slopes of 'extraversion' separately for each group. In the NL group, higher 'extraversion' was associated with a greater proportion of utilitarian responses (slope = 0.071, SE = 0.030, 95% CI [0.013, 0.129]). In contrast, in the FL group, the slope was smaller and not statistically significant (slope = 0.015, SE = 0.024, 95% CI [-0.031, 0.061]). These results indicate that the relationship between extraversion and moral decision-making differs depending on the language context, with a meaningful positive association observed only in the NL group.

Table 3B presents the characteristics of the final model, which includes all the significant predictors and the two-way interaction between 'language group' and 'extraversion', along with β -values (standardized effect sizes).

4.3. Analysis of impersonal moral dilemmas

This analysis included the experimental predictors 'language group' (native vs. foreign), and all Big Five factors, as well as the control predictors 'age' and 'educational attainment'. The binary dependent variable, 'type of decision' (utilitarian vs. deontological), included only responses to impersonal dilemmas.

No significant differences were found between the full model and any reduced model, indicating that utilitarian decisions in impersonal dilemmas were not influenced by 'language group'

($\chi^2(1) < 1$), 'age' ($\chi^2(1) < 1$), 'educational attainment' ($\chi^2(1) < 1$), 'extraversion' ($\chi^2(1) < 1$), agreeableness ($\chi^2(1) = 2.8, p = 0.094$), 'emotional stability' ($\chi^2(1) = 1.82, p = 0.177$), 'conscientiousness', ($\chi^2(1) < 1$), or 'openness' ($\chi^2(1) < 1$). Since none of the predictors showed a significant effect on the dependent variable, the final model retained no fixed effects. As a result, and in line with our reporting strategy, this model is not included in Table 3, which presents only final models with at least one retained predictor.

In summary, we replicated the well-established response patterns previously observed in traditional moral dilemmas, as well as in the FLe literature. As expected, the percentage of utilitarian choices was higher in impersonal dilemmas, moderate in high-conflict personal dilemmas, and scarce in low-conflict personal dilemmas. Figure 1 illustrates this global pattern by plotting the average utilitarianism. This average is calculated from each participant's computed utilitarianism, which is the percentage of all utilitarian decisions (i.e., 'yes' responses) across the different types of moral dilemmas. Figure 1 also shows that while both language groups exhibited this overall pattern, there was a significant increase in utilitarian responses in the FL group compared to the NL group in personal dilemmas, with no between-group differences observed in impersonal dilemmas. This consistent pattern of results provides robustness and reliability to our data. Critical to the main question of the study, we observed the influence of three personality factors after controlling for the effects of age. Figure 2 provides an overview of this influence with a frequency plot of utilitarianism as a function of each relevant personality factor. In these plots, utilitarianism represents an individual measure for each participant, computed as the percentage of all utilitarian decisions ('yes' responses) across personal moral dilemmas (whether high or low conflict). Thus, lower values in utilitarianism indicate a preference for deontological choices. Given our hypothesis that the FLe is generally modest because the effects of an FL cannot override particularly strong deontological positions influenced by certain personality factors, interpreting these results in terms of deontological preferences is more insightful. Therefore, we focus on how personality traits lead to lower scores on the y-axis

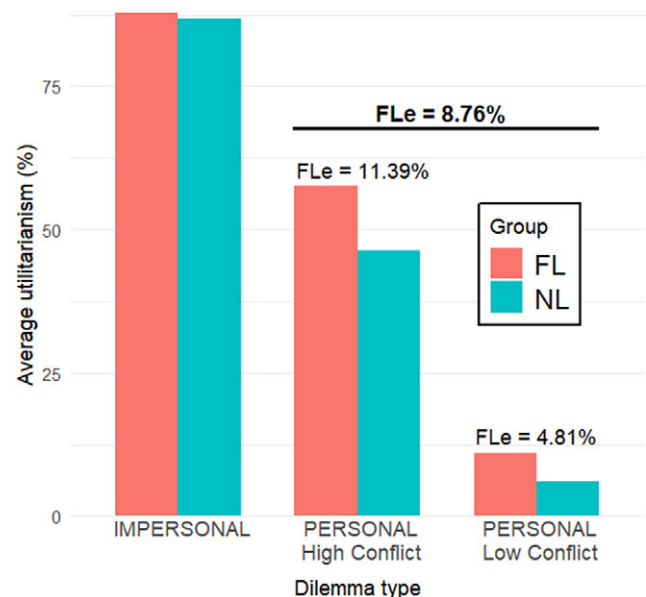


Figure 1. Average utilitarianism across dilemma type and language groups.

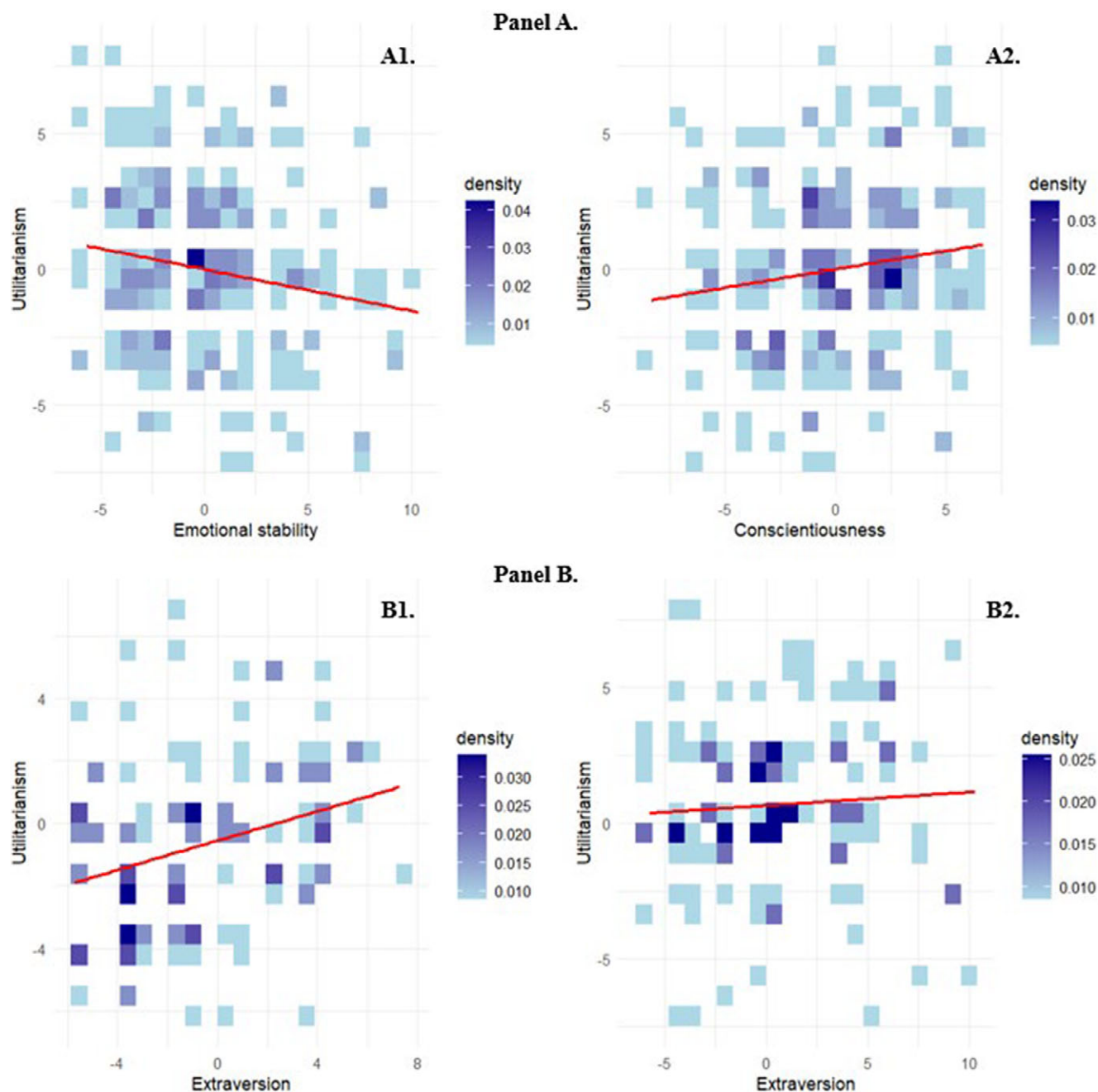


Figure 2. Influence of personality factors on moral decision-making across language groups. Panel A: Effects of emotional stability and Conscientiousness regardless of language group (A1: Emotional stability, A2: Conscientiousness). Panel B: Differential effect of extraversion by language group (B1: NL group, B2: FL group).

(utilitarianism). Panel A illustrates how two personality factors influenced decision-making in moral dilemmas, collapsing across both language groups: deontological choices were associated with higher emotional stability (Figure 2-A1) and lower conscientiousness (Figure 2-A2), irrespective of language. Panel B illustrates that deontological choices were favored by a third factor, lower extraversion, specifically in the NL group (Figure 2-B1), with no significant effect observed in the FL group (Figure 2-B2). It is important to note that, although the FL group had slightly higher extraversion values than the NL group, Figure 2-B2 illustrates an even distribution of data points across the extraversion axis, indicating a sufficient range and distribution of extraversion

scores within the FL group. Therefore, the absence of utilitarianism in the FL group cannot be attributed to a lack of lower extraversion values.

5. Discussion

In our study, we aimed to investigate whether personality factors moderate the FLe in traditional moral decision-making settings, specifically when contrasting utilitarian and deontological decisions. This objective stemmed from the observation that, although the FLe is consistently significant across various studies, its magnitude is generally small to moderate (e.g., Circi et al., 2021; Del

Maschio et al., 2022), suggesting that individual differences may modulate the effect. We hypothesized that individuals with strong personality factors favoring deontological choices would make such choices regardless of language, thereby experiencing minimal FLe. Our results support this hypothesis: individuals in the NL group with lower conscientiousness and higher emotional stability tended toward deontological choices, and this trend persisted in the FL group among individuals with similar personality factors. However, an exception to this pattern was observed with extroversion: lower extroversion predicted deontological choices in the NL group, but this influence was not evident in the FL group.

Taken together, the standardized effect sizes reported in Table 3 and the observed ΔAIC values indicate that the magnitude of these effects ranges from small to moderate, including the FLe itself, consistent with prior literature. This pattern aligns with expectations given the moderate size of the FLe and suggests that while personality factors contribute meaningfully to explaining variability in the FLe, they represent only part of a broader constellation of influences. Nonetheless, the presence of these small-to-moderate effects underscores the importance of interpreting and understanding the role of the personality traits identified in this study as modulators of the FLe.

Interpreting these findings – representing the first evidence that personality traits modulate the FLe – is challenging, partly due to the limited and sometimes inconsistent literature on personality's role in moral judgment. While Luke and Gawronski (2022b) identified conscientiousness as a predictor of deontological responding, we treated their results as a reference point rather than a strong directional hypothesis. Under our methodological conditions, we observed that lower conscientiousness, rather than higher, was associated with a greater tendency toward deontological choice. Accordingly, our interpretations remain tentative but offer a plausible account, grounded in the available evidence, of how personality traits may influence moral decision-making and, in turn, modulate the FLe. In developing our interpretations, we draw on growing body of evidence suggesting that the increase in utilitarian decisions observed in the FLe is more likely due to a reduction in deontological responses, rather than enhanced deliberation (Białek et al., 2019; Hayakawa et al., 2017; Muda et al., 2018). This shift has been attributed to blunted emotional reactions when reasoning in a foreign language, particularly reduced aversion to norm violations (Corey et al., 2017; Geipel et al., 2015b). Our results imply that this emotional diminishment occurs specifically in individuals with low, but not high, levels of extraversion, and is independent of the influence of emotional stability and conscientiousness. We acknowledge, however, that we did not directly measure emotional responses in the present study. Prior work has examined emotion-related processes in moral FLe contexts, with mixed evidence. For example, Kyriakou et al. (2023) found that participants used less emotional language in foreign language moral justifications, which partially mediated the effect of language on moral choices. In contrast, other studies (e.g., Kyriakou & Mavrou, 2023) reported no significant differences in emotional experience across languages. Therefore, all interpretations invoking emotional attenuation throughout this Discussion should be considered tentative and exploratory, intended to generate hypotheses for future research rather than to draw firm conclusions. Next, we propose a potential rationale that we believe may guide further research, based on the idea that the nature of the aversive feelings differs across these personality types.

For individuals with high emotional stability, it is plausible that the aversive feelings associated with utilitarian choices center more

intensely on the immediate victim – the person being sacrificed. This heightened sensitivity may stem from lower tendencies toward anger and emotional avoidance, dispositional features often observed in individuals with lower levels of neuroticism. Notably, the anger facet of neuroticism may reflect psychological mechanisms linked to avoidant attachment – such as emotional distancing and discomfort with caregiving. Robinson et al. (2015) suggest that avoidantly attached individuals tend to be unsettled by emotional dependence from others, due to their strong preference for autonomy and independence (Shaver et al., 2010). This caregiving discomfort is not limited to romantic relationships (Feeney & Collins, 2001) but may generalize to distressing social situations more broadly. In moral dilemmas, the sacrificed victim represents a vivid, emotionally salient individual in need – precisely the kind of case that avoidant individuals may respond to with muted empathy. Consistent with this, Robinson et al. argue that avoidant individuals exhibit greater empathy deficits when the target of harm is a specific, vividly imagined individual – such as the sacrificial victim in a typical moral dilemma – compared to when harm affects a more abstract group. Although the victim in such scenarios is hypothetical, the singular nature of the individual often elicits stronger emotional responses than group-level considerations – a difference that may account for their stronger utilitarian tendencies in such contexts (Cameron & Payne, 2011; Lickel et al., 2000; Slovic, 2007). In contrast, emotionally stable individuals – who are less prone to anger, emotional dysregulation, and avoidance – are more likely to experience spontaneous empathic concern for the immediate victim. It is possible that these emotionally grounded, deontologically aligned reactions are less easily disrupted by emotionally distancing factors like foreign language use – a possibility consistent with our finding that emotionally stable individuals were less susceptible to the FLe.

Regarding low conscientiousness, the aversive feelings may arise from concerns about violating internal moral norms. Higher conscientious individuals tend to prioritize organization, self-discipline, and orderliness – traits consistently linked to goal-directed behavior and adherence to external standards (John & Srivastava, 1999; Roberts et al., 2009). These tendencies often align with promoting the greater good through outcome-based reasoning, particularly in high-stakes contexts. In contrast, individuals lower in conscientiousness may be less bound by external rules or pragmatic goals and may instead act in accordance with internal moral convictions – especially in sacrificial moral dilemmas where harm to others is involved. It has been observed that when individuals possess a strong moral identity – that is, they view moral values as central to their self-concept (Aquino & Reed II, 2002; Black & Reynolds, 2016) – they are likely to reject utilitarian actions that involve direct harm, regardless of the consequences. Thus, an FL may not significantly reduce their emotional aversion to violating moral norms, as this aversion stems from deeply held personal values rather than from external expectations or outcome-based reasoning. This interpretation aligns with recent work in bilingual populations showing that moral identity robustly predicts deontological judgments across both first and second languages, and across different types of moral dilemmas, with little to no modulation by language context (e.g., Mavrou et al., 2025a, 2025b). Together, these findings suggest that strongly internalized moral values may exert a more powerful influence on moral decision-making than the foreign language effect itself, effectively dampening individuals against language-induced shifts in moral judgment.

Low extraverts (i.e., introverts) may experience aversive feelings linked to the utilitarian action that stem from the negative

consequences for themselves, such as feeling uncomfortable or remorseful about being responsible for sacrificing someone. This interpretation seems reasonable, as it contrasts with the well-documented tendency of high extraverts to display greater pro-social behavior, including sociability, assertiveness, and concern for others' well-being (Carlo et al., 2005; Ozer & Benet-Martínez, 2006). These traits have been linked to greater empathic engagement and helping behavior, particularly in emotionally salient situations (Habashi et al., 2016). Interestingly, according to our results, the negative feelings associated with personal repercussions for low extraverts may be modulated by the reduced emotionality afforded by using a foreign language. In this regard, it is worth noting that previous work has shown that foreign language use can attenuate self-directed moral emotions, such as guilt, in both hypothetical moral scenarios and autobiographical memories (e.g., Kyrriakou et al., 2024). A promising avenue for future research is to investigate whether foreign language use differentially modulates self-directed moral emotions, such as guilt, discomfort, or remorse, as a function of personality traits, with a particular focus on extraversion.

This study contributes to the emerging body of research exploring how individual differences interact with the FLe in moral decision-making. However, our results must be interpreted in light of certain limitations. One notable challenge arises from the inconsistency between our findings and those reported by Luke and Gawronski (2022b) regarding the role of conscientiousness in moral judgment. Although their work did not examine the FLe, their finding that higher conscientiousness predicts more deontological responding shaped our expectations about how this trait might modulate the effect. Our results, by contrast, suggest the opposite pattern: lower conscientiousness was associated with stronger deontological inclinations. We argue that this discrepancy reflects unresolved theoretical and operational issues in the field. First, there is considerable variability in the types of moral dilemmas used across studies, with some focusing on classical trolley-type scenarios and others employing more complex designs, such as those used in the CNI model. Second, most research – including ours – relies on broad personality dimensions, yet emerging evidence suggests that moral decision-making may be more strongly linked to specific lower-order facets within these dimensions. These two factors – dilemma types and personality granularity – may partly explain the conflicting findings and, more importantly, represent broader limitations in the study of personality effects on moral decision-making and, by extension, how personality may modulate the FLe. We elaborate on each of these issues below.

A further set of limitations concerns the generalizability of our findings, which we also discuss in more detail below, with special focus on the gender-homogeneous sample. We additionally consider the potential role of foreign language proficiency in shaping the strength of the FLe – an issue that extends beyond our study and remains a broader challenge for the field.

5.1. Personality granularity

The Big Five factors represent broad personality traits encompassing diverse characteristics, typically linked to stable behavioral tendencies (John et al., 2008). However, in moral decision-making, such broad dimensions may obscure important individual differences better captured at the facet level. This may explain divergences between our findings and those of Luke and Gawronski (2022b), who found differing effects of conscientiousness and null results for emotional stability and extraversion. Although both

studies measured the same broad traits, they likely tapped into different underlying facets, some promoting rule-based (deontological) responses and others goal-driven (utilitarian) reasoning. The use of different instruments – the Mini-IPIP-PW here versus the BFI-2-S in their study – supports this interpretation, as these scales do not fully overlap in facet coverage.

This argument aligns with prior evidence on personality and moral decision-making, such as that discussed by Robinson et al. (2015), who showed that specific features – like empathy deficits in avoidantly attached individuals – tend to reduce deontological responses. Importantly, such features are not necessarily unique to any one trait domain: reduced empathy and emotional detachment, for instance, can also characterize individuals high in psychopathy or Machiavellianism (Bartels & Pizarro, 2011; Djeriouat & Trémolière, 2014; Koenigs et al., 2012). This suggests that it may be more informative to focus on specific affective and interpersonal tendencies themselves – regardless of the broader personality category they fall under – when examining individual differences in moral decision-making.

Notably, likely reflecting the use of broad personality dimensions, the observed effects – though statistically significant – were small to moderate in size. It is therefore plausible that more pronounced modulatory effects on the FLe would emerge when targeting individual differences at the facet level. In line with this perspective, Gleichgerricht and Young (2013) found that lower levels of empathic concern predicted more utilitarian moral judgments, further supporting the importance of focusing on specific emotional dispositions rather than broad personality traits when investigating moral reasoning. Also relevant in this regard, features such as emotional callousness, lack of remorse, or cruelty may play a more decisive role in moral choices than broader categories like the Big Five. Supporting this view, Flexas et al. (2023) found that cruelty predicted moral responses similarly in both first and second languages. However, they did not observe an FLe, likely because their participants were early bilinguals whose second language was already emotionally grounded. To better understand how personality modulates the FLe in moral decision-making, future studies should examine these fine-grained emotional traits in populations where the effect is more robust – such as late bilinguals or speakers with limited emotional resonance in the foreign language.

5.2. Dilemma types

A second limitation concerns the nature of the moral dilemmas used and the theoretical framework they reflect. Our study employed classical sacrificial dilemmas – such as trolley-type scenarios – that contrast deontological and utilitarian choices. While effective for eliciting moral conflict, these dilemmas have been criticized for limited ecological validity and conceptual ambiguity (e.g., Bialek & De Neys, 2017; Christensen & Gomila, 2012). A key issue is that deontological choices (e.g., not pushing one person to save five) may reflect various processes – moral norm adherence, harm aversion, risk aversion, or a bias toward inaction. The CNI model (Gawronski et al., 2017; Luke & Gawronski, 2022b) was designed to disentangle these components – Consequences, Norms, and Inaction – by assessing each independently. This is particularly relevant given findings that the FLe, under the CNI framework, may involve reduced sensitivity to norms and heightened focus on outcomes (Bialek et al., 2019). Because our design does not isolate these processes, it limits our ability to determine whether the personality traits that reduced the FLe did so by reinforcing norm adherence, harm aversion, or inaction bias.

This difference in ecological validity between our dilemmas and those used by Luke and Gawronski highlights a broader issue: how personality traits shape moral decision-making across varying types of dilemmas, particularly those with greater real-world relevance. Clarifying this interaction is crucial for understanding how such traits may modulate the FLe. A promising direction would be to use more ecologically grounded scenarios, such as those employed by Geipel et al. (2015b), who found that foreign language use promoted more outcome-focused (i.e., utilitarian) reasoning in everyday moral contexts – for example, deciding whether to give money to a hungry boy who later dies of an overdose after buying drugs. It remains to be seen whether the FLe observed in such realistic scenarios is similarly moderated by personality traits like conscientiousness or emotional stability, as our findings suggest. Evidence that moral identity – an internal, norm-based disposition – exerts a strong influence toward deontological choices across both unrealistic and realistic dilemmas, independent of language (Mavrou et al., 2025a, 2025b), is consistent with this possibility.

Supporting this approach, Dewaele et al. (2024) showed that emotional intelligence and language emotionality affect judgments of real-life moral transgressions, while Woumans et al. (2020) demonstrated the FLe in applied, crime-related scenarios relevant to professional contexts. Together, these studies underscore the value of moving beyond sacrificial dilemmas to better capture how individual differences interact with linguistic context in shaping moral judgment.

Additionally, it has to be acknowledged that the use of sacrificial dilemmas such as those employed here (from the Brief Moral Dilemma battery by Carmona-Perera et al., 2015) carries certain limitations that have long been noted in the literature – for instance, inconsistent wording across dilemmas (e.g., the use of emotionally charged verbs like ‘kill’ in some scenarios and more neutral terms like ‘save’ in others), which may introduce bias in participants’ responses (Christensen et al., 2014). Beyond incidental variability in wording across items – such as the use of emotionally charged verbs like ‘kill’ or ‘sacrifice’ in some scenarios and more neutral formulations like ‘save’ in others – there are systematic differences in wording and motivational structure across dilemma types. In particular, dilemmas commonly classified as *high conflict* are often the only ones that combine several features: (i) emotionally charged language, (ii) high personal involvement, and (iii) self-beneficial outcomes, in which the agent’s own life or well-being is directly at stake. These features are known to influence moral judgments independently of the underlying moral principles involved (Christensen et al., 2014). One implication is that personality traits favoring deontological responding may exert their modulatory influence on the FLe more readily in *low conflict*, other-regarding dilemmas, whereas the strong utilitarian pull induced by self-beneficial, high-conflict dilemmas may constrain the extent to which foreign language use can further shift moral choices. Importantly, we did not observe any interaction involving dilemma type (high vs. low conflict) in the present study. This finding should be interpreted cautiously, however, as conflict level was not experimentally manipulated. Nonetheless, related work illustrates how strongly dilemma content can shape moral judgments independently of language. For example, Mills and Nicoladis (2023) showed in a bilingual sample that participants were more likely to make deontological choices in dilemmas involving saving others, but more utilitarian choices when their own lives were at stake, with no FLe observed in either case. Although their study did not examine personality traits, these findings underscore that

content-driven factors, such as self-relevance and motivational structure, can dominate moral responding across languages, thereby constraining the conditions under which the FLe and its modulation by individual differences are detectable.

A related issue is that some dilemmas share similar underlying scenarios (e.g., the classic trolley dilemma and its footbridge variant), which may reduce the conceptual distinctiveness of the items and potentially amplify or dampen emotional responses through repetition effects. Although the use of these dilemmas facilitates replicability of the FLe due to their widespread use in the literature, future research using sacrificial dilemmas may consider more methodologically robust and carefully validated alternatives, such as the battery developed by Christensen et al. (2014), available in multiple languages.

5.3. Generalizability of our findings

Although our study has limitations regarding generalizability, certain design features support broader applicability. Unlike much prior research relying predominantly on young university students (often psychology majors), our sample included adults with at least a high school education and ranged in age from 18 through late middle age – adding meaningful age diversity and enhancing ecological validity. Notably, age predicted moral decision-making: younger participants tended toward more utilitarian judgments, consistent with prior work (e.g., Mills & Nicoladis, 2023). This effect suggests that age may serve as an independent factor modulating sensitivity to moral outcomes in foreign language contexts. While age was not experimentally manipulated, these patterns indicate it could merit explicit investigation alongside personality traits in future studies.

A notable limitation for generalizability is the all-female composition of our sample. While this aligns with several foundational FLe studies (e.g., Cipolletti et al., 2016; Corey et al., 2017; Driver, 2020; Geipel et al., 2015a [studies 1 & 3], 2015b; Geipel et al., 2016; Hadjichristidis et al., 2015), it restricts broader applicability – particularly given known gender differences in personality traits. Although findings are somewhat mixed, large-scale studies consistently suggest that emotional stability (i.e., low neuroticism) tends to be higher in men than in women (Costa et al., 2001; Weisberg et al., 2011). This is relevant because our results showed that greater emotional stability predicted more deontological responses and reduced susceptibility to the FLe. On that basis, one might expect men to exhibit more deontological choices. However, this prediction conflicts with empirical evidence: a meta-analysis by Friesdorf et al. (2015) found that women, not men, tend to show stronger deontological inclinations – largely due to heightened emotional aversion to harm.

This apparent contradiction may stem from differences in the specific *facets* of emotional stability. For instance, women may report higher anxiety or vulnerability – facets less relevant to moral judgment – while simultaneously exhibiting greater empathic concern, which more directly drives deontological responding. Indeed, women consistently report and demonstrate higher empathy (Christov-Moore et al., 2014; Mestre et al., 2009), a trait closely tied to aversion to harming others. This interpretation aligns with our proposal that individuals high in emotional stability – particularly those low in anger and emotional detachment – may experience greater spontaneous empathy toward the victim, which in turn enhances emotional resistance to harm and dampens the FLe. Taken together, these considerations suggest that gender-linked traits – or specific underlying facets such as empathy and emotional

aversion – may modulate susceptibility to the FLe. Future research should examine these interactions directly.

Finally, an open question concerns the role of foreign language proficiency and experience in the FLe. Some argue the FLe diminishes with higher proficiency due to more fluent emotional processing (Caldwell-Harris, 2015; Harris, 2004; Pavlenko, 2017), while others suggest reduced cognitive effort in less proficient speakers explains this attenuation (Hayakawa et al., 2017). However, evidence remains inconclusive. In our study, self-rated proficiency, age of acquisition, and LexTALE scores did not correlate with utilitarian choices in the FL group, consistent with meta-analyses showing no clear moderating effect of proficiency (Circi et al., 2021; Del Maschio et al., 2022). Some findings hint that the FLe may be stronger among lower self-reported reading proficiency individuals (Stankovic et al., 2022), though this may reflect usage frequency or context rather than proficiency per se. More objective and nuanced measures of FL experience – including exposure, frequency, and emotionally relevant contexts – are needed. Such profiles could influence how personality traits modulate moral reasoning, an interaction that requires further study.

In conclusion, our findings suggest that certain personality traits can buffer individuals against the FLe in moral decision-making, helping to explain its typically modest magnitude. However, this resistance is not universal – some trait-driven deontological tendencies, such as those linked to low extraversion, remain susceptible to FL-induced shifts. As the first study to explore this interaction, our results provide a foundation for future research to develop more targeted hypotheses. They also highlight the need for refined trait-level analyses, improved dilemma selection, and attention to sample characteristics such as gender and personality profiles. Understanding how emotional engagement varies across personality types and language contexts will be crucial in further clarifying individual variability in the FLe.

Data availability statement. The datasets generated during the current study are available from the corresponding author on reasonable request.

Funding statement. MH was supported by grant PID2021-127053NB-I00 funded by MICIU/AEI/10.13039/501100011033/FEDER, UE, and grant 2021 SGR 01099 (AGAUR). JRF was supported by grant PID2022-138016NB-I00, funded by MICIU/AEI/10.13039/501100011033, FEDER/UE, and grant 2021 SGR 01102 (AGAUR).

Competing interests. The authors declare no conflict of interest.

References

- Abbasi-Asl, R., & Hashemi, S. (2019). Personality and morality: Role of the big five personality traits in predicting the four components of moral decision making. *International Journal of Behavioral Sciences*, 13(3), 123–128.
- Aquino, K., & Reed II, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, 83(6), 1423. <https://doi.org/10.1037/0022-3514.83.6.1423>.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>.
- Bartels, D. M., & Pizarro, D. A. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition*, 121(1), 154–161. <https://doi.org/10.1016/j.cognition.2011.05.010>.
- Bialek, M., & De Neys, W. (2017). Dual processes and moral conflict: Evidence for deontological reasoners' intuitive utilitarian sensitivity. *Journal of Cognitive Psychology*, 29(5), 559–574. <https://doi.org/10.1080/20445911.2016.1156118>.
- Bialek, M., Paruzel-Czachura, M., & Gawronski, B. (2019). Foreign language effects on moral dilemma judgments: An analysis using the CNI model. *Journal of Experimental Social Psychology*, 85, 103855. <https://doi.org/10.1016/j.jesp.2019.103855>.
- Black, J. E., & Reynolds, W. M. (2016). Development, reliability, and validity of the moral identity questionnaire. *Personality and Individual Differences*, 97, 120–129. <https://doi.org/10.1016/j.paid.2016.03.041>.
- Braida, B., Rodríguez-Ferreiro, J., & Hernández, M. (2023). The foreign language effect on motivational quotes. *Bilingualism: Language and Cognition*, 26(2), 416–424. <https://doi.org/10.1017/S1366728922000505>.
- Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed.). Springer.
- Caldwell-Harris, C. L. (2015). Emotionality differences between a native and foreign language: Implications for everyday life. *Current Directions in Psychological Science*, 24(3), 214–219. <https://doi.org/10.1177/0963721414566268>.
- Cameron, C. D., & Payne, B. K. (2011). Escaping affect: How motivated emotion regulation creates insensitivity to mass suffering. *Journal of Personality and Social Psychology*, 100(1), 1–15. <https://doi.org/10.1037/a0021643>.
- Carlo, G., Okun, M. A., Knight, G. P., & de Guzman, M. R. T. (2005). The interplay of traits and motives on volunteering: Agreeableness, extraversion, and prosocial value motivation. *Personality and Individual Differences*, 38(6), 1293–1305. <https://doi.org/10.1016/j.paid.2004.08.012>.
- Carmona-Perera, M., Caracul, A., Pérez-García, M., & Verdejo-García, A. (2015). Brief moral decision-making questionnaire: A Rasch-derived short form of the Greene dilemmas. *Psychological Assessment*, 27(2), 424. <https://doi.org/10.1037/pas0000049>.
- Christensen, J. F., Flexas, A., Calabrese, M., Gut, N. K., & Gomila, A. (2014). Moral judgment reloaded: A moral dilemma validation study. *Frontiers in Psychology*, 5, 607. <https://doi.org/10.3389/fpsyg.2014.00607>.
- Christensen, J. F., & Gomila, A. (2012). Moral dilemmas in cognitive neuroscience of moral decision-making: A principled review. *Neuroscience and Biobehavioral Reviews*, 36(4), 1249–1264. <https://doi.org/10.1016/j.neubiorev.2012.02.008>.
- Christov-Moore, L., Simpson, E. A., Coudé, G., Grigaitytė, K., Iacoboni, M., & Ferrari, P. F. (2014). Empathy: Gender effects in brain and behavior. *Neuroscience & Biobehavioral Reviews*, 46, 604–627. <https://doi.org/10.1016/j.neubiorev.2014.09.001>.
- Cipolletti, H., McFarlane, S., & Weissglass, C. (2016). The moral foreign-language effect. *Philosophical Psychology*, 29(1), 23–40. <https://doi.org/10.1080/09515089.2014.993063>.
- Circi, R., Gatti, D., Russo, V., & Vecchi, T. (2021). The foreign language effect on decision-making: A meta-analysis. *Psychonomic Bulletin & Review*, 28, 1–11. <https://doi.org/10.3758/s13423-020-01871-z>.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Conway, P., Goldstein-Greenwood, J., Polacek, D., & Greene, J. D. (2018). Sacrificial utilitarian judgments do reflect concern for the greater good: Clarification via process dissociation and the judgments of philosophers. *Cognition*, 179, 241–265. <https://doi.org/10.1016/j.cognition.2018.04.018>.
- Corey, J. D., Hayakawa, S., Foucart, A., Aparici, M., Botella, J., Costa, A., & Keysar, B. (2017). Our moral choices are foreign to us. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(7), 1109. <https://doi.org/10.1037/xlm0000356>.
- Costa, A., Duñabeitia, J. A., & Keysar, B. (Eds.). (2019). *Language, emotion and decision making* [special issue]. *Quarterly Journal of Experimental Psychology*, 72(1).
- 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. <https://doi.org/10.1371/journal.pone.0094842>.
- Costa, P. T., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, 81(2), 322–331. <https://doi.org/10.1037/0022-3514.81.2.322>.
- Del Maschio, N., Crespi, F., Peressotti, F., Abutalebi, J., & Sulpizio, S. (2022). Decision-making depends on language: A meta-analysis of the foreign language effect. *Bilingualism: Language and Cognition*, 25(4), 617–630. <https://doi.org/10.1017/S1366728921001012>.
- Dewaele, J. M., Mavrou, I., Kyriakou, A., & Lorette, P. (2024). The role of language and emotional intelligence in judgments of real-life social and moral

- transgressions among Greek, Hungarian, and British users of English. *Journal of Behavioral Decision Making*, 37(1), e2356. <https://doi.org/10.1002/bdm.2356>.
- Djeriouat, H., & Trémolière, B. (2014). The dark triad of personality and utilitarian moral judgment: The mediating role of honesty/humility and harm/care. *Personality and Individual Differences*, 67, 11–16. <https://doi.org/10.1016/j.paid.2013.12.026>.
- Driver, M. Y. (2020). Switching codes and shifting morals: how code-switching and emotion affect moral judgment. *International Journal of Bilingual Education and Bilingualism*, 1–17. <https://doi.org/10.1080/13670050.2020.1730763>
- Feeney, B. C., & Collins, N. L. (2001). Predictors of caregiving in adult intimate relationships: An attachment theoretical perspective. *Journal of Personality and Social Psychology*, 80(6), 972–994. <https://doi.org/10.1037/0022-3514.80.6.972>.
- Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. London, UK: Sage Publications.
- Flexas, A., López-Penadés, R., Aguilar-Mediavilla, E., & Adrover-Roig, D. (2023). Meanness trumps language: Lack of foreign language effect in early bilinguals' moral choices. *PLoS One*, 18(11), e0294523. <https://doi.org/10.1371/journal.pone.0294523>.
- Foot, P. (1967). *The problem of abortion and the doctrine of double effect* (Vol. 5, pp. 5–15). Oxford.
- Friesdorf, R., Conway, P., & Gawronski, B. (2015). Gender differences in responses to moral dilemmas: A process dissociation analysis. *Personality and Social Psychology Bulletin*, 41(5), 696–713. <https://doi.org/10.1177/0146167215575731>.
- Gawronski, B., Armstrong, J., Conway, P., Friesdorf, R., & Hütter, M. (2017). Consequences, norms, and generalized inaction in moral dilemmas: The CNI model of moral decision-making. *Journal of Personality and Social Psychology*, 113(3), 343. <https://doi.org/10.1037/pspa0000086>.
- Geipel, J., Hadjichristidis, C., & Surian, L. (2015a). The foreign language effect on moral judgment: The role of emotions and norms. *PLoS One*, 10(7), e0131529. <https://doi.org/10.1371/journal.pone.0131529>.
- Geipel, J., Hadjichristidis, C., & Surian, L. (2015b). How foreign language shapes moral judgment. *Journal of Experimental Social Psychology*, 59, 8–17. <https://doi.org/10.1016/j.jesp.2015.02.001>.
- Geipel, J., Hadjichristidis, C., & Surian, L. (2016). Foreign language affects the contribution of intentions and outcomes to moral judgment. *Cognition*, 154, 34–39. <https://doi.org/10.1016/j.cognition.2016.05.010>.
- Gleichgerrcht, E., & Young, L. (2013). Low levels of empathic concern predict utilitarian moral judgment. *PLoS One*, 8(4), e60418. <https://doi.org/10.1371/journal.pone.0060418>.
- Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, 107(3), 1144–1154. <https://doi.org/10.1016/j.cognition.2007.11.004>.
- Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, 44(2), 389–400. <https://doi.org/10.1016/j.neuron.2004.09.027>.
- Habashi, M. M., Graziano, W. G., & Hoover, A. E. (2016). Searching for the prosocial personality: A big five approach to linking personality and prosocial behavior. *Personality and Social Psychology Bulletin*, 42(9), 1177–1192. <https://doi.org/10.1177/0146167216652859>.
- Hadjichristidis, C., Geipel, J., & Savadori, L. (2015). The effect of foreign language in judgments of risk and benefit: The role of affect. *Journal of Experimental Psychology: Applied*, 21(2), 117–129. <https://doi.org/10.1037/xap0000044>.
- Harris, C. L. (2004). Bilingual speakers in the lab: Psychophysiological measures of emotional reactivity. *Journal of Multilingual and Multicultural Development*, 25(2–3), 223–247. <https://doi.org/10.1080/01434630408666530>.
- 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(10), 1387–1397. <https://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(3), 489. <https://doi.org/10.1037/xlm0000179>.
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative big five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114–158). The Guilford Press.
- John, O. P., & Srivastava, S. (1999). The big five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–138). Guilford Press.
- Khan, T., Akbar, A., Jam, F., & Saeed, M. (2016). A time-lagged study of the relationship between big five personality and ethical ideology. *Ethics & Behavior*, 26, 488–506. <https://doi.org/10.1080/10508422.2015.1055493>.
- Koenigs, M., Kruepke, M., Zeier, J., & Newman, J. P. (2012). Utilitarian moral judgment in psychopathy. *Social Cognitive and Affective Neuroscience*, 7(6), 708–714. <https://doi.org/10.1093/scan/nsr048>.
- Kyriakou, A., Foucart, A., & Mavrou, I. (2023). Moral judgements in a foreign language: Expressing emotions and justifying decisions. *International Journal of Bilingualism*, 27(6), 978–995. <https://doi.org/10.1177/13670069221134193>.
- Kyriakou, A., & Mavrou, I. (2023). What language does your heart speak? The influence of foreign language on moral judgements and emotions related to unrealistic and realistic moral dilemmas. *Cognition & Emotion*, 37(8), 1–19. Advance online publication. <https://doi.org/10.1080/02699931.2023.2258577>.
- Kyriakou, A., Mavrou, I., & Palapanidi, K. (2024). The role of foreign language in the experience and emotional expression of guilt: Evidence from moral scenarios and autobiographical memories of bilinguals. *International Journal of Bilingual Education and Bilingualism*, 27(10), 1407–1421. <https://doi.org/10.1080/13670050.2024.2365238>.
- Lange, R. T., Iverson, G. L., Brooks, B. L., & Ashton Rensson, V. L. (2010). Influence of poor effort on self-reported symptoms and neurocognitive test performance following mild traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, 32(9), 961–972. <https://doi.org/10.1080/13803391003645657>.
- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, 44, 325–343. <https://doi.org/10.3758/s13428-011-0146-0>.
- Lickel, B., Hamilton, D. L., Wierzchowska, G., Lewis, A., Sherman, S. J., & Uhles, A. N. (2000). Varieties of groups and the perception of group tentativeness. *Journal of Personality and Social Psychology*, 78(2), 223–246. <https://doi.org/10.1037/0022-3514.78.2.223>.
- Luke, D. M., & Gawronski, B. (2022a). Big five personality traits and moral-dilemma judgments: Two preregistered studies using the CNI model. *Journal of Research in Personality*, 101, 104297. <https://doi.org/10.1016/j.jrp.2022.104297>.
- Luke, D. M., & Gawronski, B. (2022b). Temporal stability of moral dilemma judgments: A longitudinal analysis using the CNI model. *Personality and Social Psychology Bulletin*, 48(8), 1191–1203. <https://doi.org/10.1177/01461672211035024>.
- Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods*, 49(4), 1494–1502. <https://doi.org/10.3758/s13428-016-0809-y>.
- Martínez-Molina, A., & Arias, V. B. (2018). Balanced and positively worded personality short-forms: Mini-IPIP validity and cross-cultural invariance. *PeerJ*, 6, e5542. <https://doi.org/10.7717/peerj.5542>.
- Mavrou, I., Mavrou, S., & Kyriakou, A. (2025a). *Moral identity robustly determines moral judgements irrespective of the language we use*. *OSF Preprints*. https://doi.org/10.31219/osf.io/a2cb7_v1.
- Rong, X., Mavrou, I., Révész, A., Lam, S., & Zeng, T. (2025b). *Moral identity and moral arguments: Moving beyond the yes/no paradigm in bilinguals' moral decision-making research*. *OSF Preprints*. https://osf.io/preprints/osf/2svzr_v1.
- Mestre, M. V., Samper, P., Frias, M. D., & Tur, A. M. (2009). Are women more empathetic than men? A longitudinal study in adolescence. *The Spanish Journal of Psychology*, 12(1), 76–83. <https://doi.org/10.1017/S1138741600001499>.
- Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, 110(1), 40–48. <https://doi.org/10.1037/0021-843X.110.1.40>.
- Mills, S., & Nicoladis, E. (2023). It's easier to kill a baby to save oneself than a fat man to save other people: The effect of moral dilemma and age on Russian-English bilinguals' moral reasoning. *Journal of Multilingual and Multicultural Development*, 44(4), 265–274. <https://doi.org/10.1080/01434632.2020.1813145>.
- Muda, R., Niszczoła, P., Bialek, M., & 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**(2), 321. <https://doi.org/10.1037/xlm0000447>.
- Muda, R., Pieńkosz, D., Francis, K. B., & Bialek, M. (2020). The moral foreign language effect is stable across presentation modalities. *Quarterly Journal of Experimental Psychology*, **73**(11), 1930–1938. <https://doi.org/10.1177/1747021820935072>.
- Ozer, D. J., & Benet-Martínez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, **57**, 401–421. <https://doi.org/10.1146/annurev.psych.57.102904.190127>.
- Pavlenko, A. (2017). Do you wish to waive your rights? Affect and decision-making in multilingual speakers. *Current Opinion in Psychology*, **17**, 74–78. <https://doi.org/10.1016/j.copsyc.2017.06.005>.
- Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Harcourt Brace.
- R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing. <https://www.r-project.org/>.
- Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's advanced progressive matrices* (1998 ed.). Oxford Psychologists Press.
- Rengifo, M., & Laham, S. (2022). Big five personality predictors of moral disengagement: A comprehensive aspect-level approach. *Personality and Individual Differences*, **184**, 111176. <https://doi.org/10.1016/j.paid.2021.111176>.
- Roberts, B. W., Jackson, J. J., Fayard, J. V., Edmonds, G., & Meints, J. (2009). Conscientiousness. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 369–381). Guilford Press.
- Robinson, J. S., Joel, S., & Plaks, J. E. (2015). Empathy for the group versus indifference toward the victim: Effects of anxious and avoidant attachment on moral judgment. *Journal of Experimental Social Psychology*, **56**, 139–152. <https://doi.org/10.1016/j.jesp.2014.09.017>.
- Shaver, P. R., Mikulincer, M., & Shemesh-Iron, M. (2010). A behavioral systems perspective on prosocial behavior. In M. Mikulincer & P. R. Shaver (Eds.), *Prosocial motives, emotions, and behavior: The better angels of our nature* (pp. 71–87). American Psychological Association.
- Shin, H. I., & Kim, J. (2017). Foreign language effect and psychological distance. *Journal of Psycholinguistic Research*, **46**, 1339–1352. <https://doi.org/10.1007/s10936-017-9498-7>.
- Sirota, M., & Juanchich, M. (2018). Effect of response format on cognitive reflection: Validating a two-and four-option multiple choice question version of the cognitive reflection test. *Behavior Research Methods*, **50**, 2511–2522. <https://doi.org/10.3758/s13428-018-1029-4>.
- Slovic, P. (2007). “if I look at the mass I will never act”: Psychic numbing and genocide. *Judgment and Decision making*, **2**(2), 79–95. <https://doi.org/10.1017/S1930297500000061>.
- Stankovic, M., Biedermann, B., & Hamamura, T. (2022). Not all bilinguals are the same: A meta-analysis of the moral foreign language effect. *Brain and Language*, **227**, 105082. <https://doi.org/10.1016/j.bandl.2022.105082>.
- Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.
- Thomson, J. J. (1976) Killing, letting die, and the trolley problem. In *Death, dying and the ending of life*, I and II (pp. V2_17–V2_30). Routledge.
- Weisberg, Y. J., DeYoung, C. G., & Hirsh, J. B. (2011). Gender differences in personality across the ten aspects of the big five. *Frontiers in Psychology*, **2**, 178. <https://doi.org/10.3389/fpsyg.2011.00178>.
- West, B. T., Welch, K. B., & Galecki, A. T. (2015). *Linear mixed models: A practical guide using statistical software* (2nd ed.). CRC Press. <https://doi.org/10.1201/b19642>.
- Woumans, E., Van der Cruyssen, I., & Duyck, W. (2020). Crime and punishment: Morality judgment in a foreign language. *Journal of Experimental Psychology: General*, **149**(8), 1597–1602. <https://doi.org/10.1037/xge0000736>.
- Yavuz, M., Küntay, A. C., Brouwer, S. (2024). The effect of foreign language and psychological distance on moral judgment in Turkish–English bilinguals. *Bilingualism: Language and Cognition*, **27**(3), 447–459. <https://doi.org/10.1017/S1366728923000706>.