

SOME CHALLENGES FOR RESEARCH ON EMOTION AND MORAL JUDGMENT: THE MORAL FOREIGN-LANGUAGE EFFECT AS A CASE STUDY

– Steven McFarlane –
– Heather Cicolletti Perez –

Abstract: In this article, we discuss a number of challenges with the empirical study of emotion and its relation to moral judgment. We examine a case study involving the moral foreign-language effect, according to which people show an increased utilitarian response tendency in moral dilemmas when using their non-native language. One important proposed explanation for this effect is that using one’s non-native language reduces emotional arousal, and that reduced emotion is responsible for this tendency. We offer reasons to think that there is insufficient evidence for accepting this explanation at present. We argue that there are three themes that constrain our current ability to draw firm empirical conclusions: 1) the frequent use of proxies or partial measures for emotions, 2) the lack of a predictive and generalizable theory of emotion and specific emotion-types, and 3) the obscurity of a baseline level of neutrality with respect to participant emotion. These lessons apply not only to research on the moral foreign-language effect, but to empirical research in moral psychology more generally.

Keywords: moral judgment, moral psychology, emotion, foreign-language effect, FLE, moral foreign-language effect, MFLE

Published online: 12 May 2020

1. Introduction

In this paper, we illustrate some of the challenges with empirically investigating the influence of emotion on moral judgment. To focus discussion, we use a case study known as the moral foreign-language effect (or MFLE). Researchers have suggested that people who consider a moral dilemma using their non-native language respond differently than people consid-

✉ Steven McFarlane
University of Minnesota, Morris
Department of Humanities
600 East 4th Street
Morris, MN 56267
email: smcfarla@morris.umn.edu
Heather Cicolletti Perez
Broward College
Broward College Online Campus
6400 NW 6th Way
Fort Lauderdale, FL 33309
email: Hperez3@broward.edu

ering the same dilemma in their native language, and that this difference might be due to a difference in emotional processing when using one's non-native language.¹ After explaining the MFLE and its relationship to the original foreign-language effect (or FLE) more fully, we articulate challenges to the prevailing explanation that the effect is caused by reduced emotion, and argue that these challenges apply more generally to the study of emotion and moral judgment. Many of these difficulties will perhaps be overcome in the near future, but for now, we argue that a certain agnosticism is the most warranted epistemic position.

2. The Moral Foreign Language Effect

Boaz Keysar, Sayuri Hayakawa, and Sun Gyu An (2012) found that thinking in a non-native language, L2, can influence the content of our judgments and decisions relative to making decisions in one's native language, L1. The original foreign-language effect (or FLE) focused on heuristics and biases (e.g. loss aversion, the framing effect) and found that individuals reasoning in L2 were less likely to fall prey to reasoning errors than individuals reasoning in L1. Subsequent research has demonstrated that the FLE extends to moral judgments.² When participants consider moral dilemmas in L2, they are more likely to judge that it is acceptable to harm an innocent person if it saves the lives of others (labeled the utilitarian option), rather than judging that it is impermissible to harm someone for the sake of others (the deontological option).³ This particular kind of FLE, has been called the moral foreign-language effect (or MFLE).

This effect is of great interest as it suggests the same moral content can be evaluated in different ways based on how individuals cognize the content. This indicates that *how* we process a decision or moral dilemma can have some degree of influence independent from *what* we are thinking about.

Interest in the FLE and the MFLE reaches beyond academic journals. Articles about this effect – that learning another language can make us “more logical” or that we make different moral judgments due to being less emotional – have appeared in sources as diverse as *The Economist*, *NPR*, and *Travel & Leisure*. According to *Travel & Leisure*, “In addition to making a person more attractive, speaking another language may make for a more logical, utilitarian decision-maker.”⁴ According to *Bloomberg Business*, “to avoid stupid mistakes, think in French.”⁵ The wide-ranging reporting on this topic seems

¹ Cipolletti, McFarlane, Weissglass (2016); Corey, Hayakawa, Foucart et al. (2017); Costa, Foucart, Hayakawa et al. (2014); Geipel, Hadjichristidis, Surian (2015).

² Ibidem.

³ Typical examples include trolley cases involving pulling a switch that will kill one to save five (an impersonal harm) versus pushing someone off a footbridge to save five (a personal harm). People seem to have inconsistent judgments when considering cases that some consider to have parallel ethical properties. The MFLE found that individuals were more likely to choose the option commonly associated with utilitarianism in the footbridge dilemma while considering the case in L2. Some recent research suggests that this change in responses is not due to increased utilitarian judgment but to something else (e.g. being less responsive to deontological norms). We will refer to this pattern as an increase in utilitarian responses or increased utilitarian response tendencies as shorthand. We will return to some of these complexities later in the essay.

⁴ Lieberman (2017).

⁵ Drake (2012).

to serve as prescriptive advice to the general public as well, for example, if “thinking in a foreign language helps us disconnect from these emotions and make decision in a more economically rational way,”⁶ one might think that it would be *better* for us to make decisions in our non-native language or that we should trust the judgment of bilinguals speaking in a less familiar language than those who are speaking in their native tongue. If people really are tempted to give such advice, it’s important that we really understand what the FLE and MFLE truly are and how they work.

The primary explanation for the FLE offered so far is that participants experience emotional “distancing” when using their second language.⁷ This hypothesis, or the *reduced emotionality account*, is explained as follows: “a[n L2] context reduces the emotional reactivity that certain scenarios may elicit. This reduction in emotional reactivity may then impact people’s decisions, sometimes reducing biases that are related with such emotional processing (e.g., loss aversion).”⁸ Hayakawa and colleagues embrace a similar explanation which they call the *blunted-deontology* account (for reasons we explain later) that “holds that foreign-language use affects moral decisions by stunting the emotional or heuristic processing.”⁹

The reduced emotionality account is consistent with qualitative evidence that bilinguals or multilinguals consider their native language to be more emotional (and prefer to use it for expressing emotions) than those that they learned in the classroom,¹⁰ as well as experiments showing that Skin Conductance Responses (SCRs) are higher when emotional phrases or words were presented in L1 than in L2, which indicates that one is more emotionally aroused in one’s native language.¹¹

Even if the reduced emotionality account of the FLE is correct, however, we still need to explain this connection to higher frequency of utilitarian responses to moral dilemmas in the case of the MFLE. To fully understand the inference, we need to say a bit about Joshua Greene’s influential dual-process account of moral judgment.¹² Greene argues that we assess moral dilemmas using distinct cognitive processes. One type of processing, sometimes labeled Type 1 or System 1 processing, is characteristically fast, automatic, and tied to emotion. The other type of processing, Type 2 or System 2 processing, is characteristically slow, voluntary, and in Greene’s view, more rational in that it is less susceptible to emotional distortion.¹³

⁶ Gross (2012).

⁷ Keysar, Hayakawa, An (2012).

⁸ Vives, Aparici, Costa (2018): 2.

⁹ Hayakawa, Tannebaum, Costa et al. (2017): 1388. Hayakawa was a member of the original FLE research group, and it is clear that she and co-authors view the MFLE as part and parcel with the FLE. Geipel, Hadjichristidis, Surian (2015) claimed that emotion is not an explanatory variable for the MFLE, but Geipel, Hadjichristidis, Surian (2016) argues that emotions are an integral part of the explanation.

¹⁰ Dewaele (2010).

¹¹ Harris, Gleason, Ayçiçeği (2003; 2006).

¹² Greene, Sommerville, Nystrom et al. (2001); Greene (2008; 2013).

¹³ Greene’s dual-process account is influential and is cited by proponents of the emotion account of the FLE; however, there are other dual-process theories that do not associate Type 2 processing with emotion, nor do they associate Type 2 processing with utilitarian responses. See Białek, De Neys (2016; 2017) and Bago, De Neys (2018) for more. We do not discuss alternative models of dual-process theories here, as we are focused more on the relation between emotion and moral judgment, but they deserve further consideration.

Greene's research on moral dilemmas heavily relies on neuroscientific evidence. In his cases, when participants arrived at deontological responses, fMRI imaging showed "greater activity in three emotion-related areas: the posterior cingulate cortex, the medial prefrontal cortex, and the amygdala."¹⁴ Combining this evidence with evidence regarding response times and other measures, Greene argues that the connection between deontological responses and increased activity in emotional-related areas in the brain shows that emotional activity causes people to provide deontological responses. And because emotional response is one of the characteristics typically associated with Type 1 processing, he concludes that respondents are using Type 1 processing in evaluating the dilemmas. So, in his view, we have a triangulation between emotional processing, Type 1 processing, and deontological responding.¹⁵

We are now in a position to better understand the reduced emotionality account of the MFLE. Supposing Greene is correct, then a reduction in emotion should lead to less deontological responding. And that appears to be the case. So, in the context of some evidence that

1. users of L2 are less emotional,
2. deontology is tied to emotion, and that
3. users of L2 are less prone to provide deontological responses,

it might seem natural to conclude that the reduced emotionality account is correct. This type of thought is what motivates Hayakawa et al. to write: "experiments now provide an explanation [for greater frequency of utilitarian responses]. People are more utilitarian when using a foreign language not because they think more, but because they feel less."¹⁶

We argue for agnosticism regarding the role of emotion in causing the MFLE and the FLE more generally. Some research suggests that the MFLE does not hold up when participants consider several moral dilemmas taken together as a battery of cases. Hayakawa et al. (2017) find that "the MFLE was not replicated when we examined responses to the kind of moral dilemmas traditionally used to measure utilitarian responding."¹⁷ Białek, Paruzel-Czachura, & Gawronski (2019) and Muda et al. (2018) similarly report that using L2 does not affect moral judgments in a series of moral dilemmas to any significance. If this right, then our case for agnosticism is strengthened. However, the MFLE has been replicated a number of times for at least some moral dilemmas, in particular "Switch" versus "Footbridge" dilemmas, and we pursue the question of whether reduced emotion has been experimentally demonstrated as the correct explanation for this effect.

We think there is much to say on behalf of the reduced emotionality account of the MFLE, and we think it may well be the strongest explanation we have at the moment. However, our purpose here is to raise underappreciated challenges to empirically measuring the influence of emotion on moral judgment. In the spirit of advancing moral psychology, we urge continued caution toward concluding that emotion is responsible

¹⁴ Greene (2008): 44.

¹⁵ Greene's model has been challenged on many fronts. See especially Kahane, Shackel (2008), Kahane, Shackel (2010) and Kahane, Wiech, Shackel et al. (2011).

¹⁶ Hayakawa, Tannebaum, Costa et al. (2017): 1396. Geipel, Hadjichristidis, Surian (2016) also support an account that L2 "attenuates action-based affect" (37). These papers all cite Greene's work.

¹⁷ Hayakawa, Tannebaum, Costa et al. (2017): 1395.

for the MFLE, with some general lessons for the empirical study of emotion and moral judgment more generally.

We discuss three important themes in the research on emotion and moral judgment that constrain our current ability to draw firm empirical conclusions. They are: 1) the frequent use of proxies or partial measures for emotions, 2) the lack of a predictive and generalizable theory of emotion and specific emotion-types, and 3) the obscurity of a baseline level of neutrality with respect to participant emotion. Underlying our concerns is the suspicion that research groups, both across disciplines and within disciplines, import implicit sets of assumptions about the nature of emotion and how to test it into their research, and there are few universally agreed upon assumptions among these groups.

3. Proxies and Associations

Without delving too deeply into metaphysical questions, it appears that emotion is pervasive, but in some ways unobservable to familiar empirical measuring techniques, at least in the sense that there is no “emotion-ometer” that can measure the qualitative category of an emotion, nor its magnitude. Researchers tend to focus on one or more of the following five methods (we cite mere samples of voluminous literatures in each case):

1. Measurable biological responses, such as changes in skin conductance¹⁸
2. Neuroscientific evidence¹⁹
3. Behavioral response patterns²⁰
4. Patterns in judgment responses²¹
5. Qualitative reports of felt emotion²²

Sometimes multiple types of measures are used in the same experiments, though oftentimes only one type of measure is used in individual studies. There are difficulties about reconciling these measures. Suppose the emotional centers in a study participant’s brain lights up, but they do not self-report heightened emotion? This would not conclusively show that emotions were not aroused, as perhaps the emotions are not accessible to introspective self-report. Then again, we ought not to conclude that activation in the emotion centers itself is enough to attribute behavior or judgment changes to emotion either.

This means that positing emotions is abductive, and dependent on using observed effects rather than the thing itself. For instance, Greene’s research emphasizes that emotional regions of the brain “light up” when considering certain moral dilemmas. He concludes that emotional responses are responsible for the frequency of deontological judgments in these dilemmas. He might be right, but an association between a region of the brain lighting up – regardless of whether it is a region frequently associated with emotion or not – and a type of response on a survey of moral attitudes is using a proxy.

In the case of the moral foreign-language effect (or MFLE), the use of a proxy can be illustrated with the following passage:

¹⁸ Caldwell-Harris, Ayçiçeği-Dinn (2009; 2006); Harris, Ayçiçeği, Gleason (2003).

¹⁹ Greene, Sommerville, Nystrom et al. (2001); Sevinc, Spreng (2014).

²⁰ Bond, Lai (1986); Colbeck, Bowers (2012); Valdesolo, DeSteno (2006).

²¹ Greene, Sommerville, Nystrom et al. (2001); Haidt (2001); Hayakawa, Tannebaum, Costa et al. (2017).

²² Dewaele (2010); Harris, Ayçiçeği, Gleason (2003); Horne, Powell (2016).

[U]se of a foreign language might stunt emotional processing, attenuating considerations of deontological rules, such as the prohibition against killing. Across six experiments using different bilingual populations, elicitation formats and moral dilemmas, we found clear evidence that foreign-language use blunts deontological responding.²³

Note that these experiments did not test for emotion directly. Rather, the researchers take the blunting of deontological responding to be the *evidence* that use of a foreign language might be stunting emotional processing (hence the naming convention of the “blunted-deontology account”).

As a general rule, measuring proxies or associations is non-ideal – at best a second-best option – to measuring the real thing, for reasons that are readily apparent. Proxies inherently allow for greater statistical noise, hidden variables, and simply can be better or worse at capturing the relevant phenomenon based on the strength and nature of the correlations.

The matter is made worse in the case of emotion and moral judgment. Emotion itself is not very well understood, and both emotion and moral judgment are labile, sensitive to cultural values, and change over time. Additionally, the *associations* between them admit of three gradient of durations - short-term interventions of emotional arousal states, longer term affiliations with moods, and perhaps individual differences in trait-like dispositions over a lifetime. We do not have a firm grasp of either proxy as an anchoring point from which to draw inferences.²⁴

For the MFLE, the evidentiary strength of the proxy relation needs to be pretty strong in *both* directions to avoid the problem of distorted data, as the proxy relation goes both ways. In Greene’s work, arousal in specific brain regions is the proxy for emotion, and this proxy is connected with deontological judgments. In Hayakawa et al.’s work, reduced deontological responses are taken to be the proxy for a reduction of emotion.

Let us assume with Greene that it is true that we can study which brain regions regularly correlate with emotional arousal and detect when they are activated by reading moral dilemmas.

And, let us assume, that research has established that deontological responses are always perfectly correlated with emotional arousal and utilitarian responses are always correlated with a lack of this arousal in a range of moral dilemmas, not just the trolley cases. In that case, maybe it would make sense to conclude, in line with the blunted-deontology account, that a reduction of deontological responses involves a stunting of emotional processing.

However, we think that it is important to take a step back and look at the whole picture of what these hypothetical facts would show. *By itself*, activation of a brain region reg-

²³ Hayakawa, Tannebaum, Costa et al. (2017): 1394.

²⁴ One might challenge elements of this depiction of moral judgments (it is not easy to see how to challenge them when it comes to emotions). Perhaps moral judgments are more stubborn and are not subject to change due to subtle psychological or cultural influences. If that is the case, so much the worse for one common form of empirical moral psychology, including Greene’s dual-process theory and the reduced emotionality account, as these theories essentially depend upon short-term experimental manipulations.

ularly correlating with emotional arousal is not sufficient to say that the brain region *causes* emotional arousal or vice versa (or that they are to be identified). For instance, there could be one or more unidentified common causes for both brain activity and emotional arousal. But even if the correlation *were* due to causation, this hypothetical fact would not, *by itself*, be sufficient to say that this brain region activating in certain kinds of moral dilemmas causes the pattern of deontological responses. Again, there may be common causes responsible. And even if this pattern of brain activity did cause the pattern of deontological responses, it is still not the case that emotional activation would always and only be the cause of (or identical or reducible to) deontological (as opposed to utilitarian) responses. There could be common causes for brain activity, emotional response, and deontological response (though we admit the plausible candidates would be remote possibilities at this point).²⁵

Hence, these hypothetical facts are not *by themselves* sufficient to say that a blunting in deontological responses is caused by a stunting of emotional processing. The only way these inferences would be justified is if we are either simply identifying deontological responding with a type of emotional processing (which Greene does)²⁶ or assuming an unverified perfect correlation. In the absence of such strong relations, there is room for multiple explanations or spurious correlations at each stage, or of predictable variance. And this has all been assuming perfect correlations, which is far from the case in the actual world. Given that the correlations are imperfect, the case for using emotion as proxy for deontological responses *and* utilitarian responses as a proxy for reduced emotionality is even weaker.

Again, our intention is not to argue that these views are incorrect, or that there is anything wrong with the research at each stage of this process, *per se*. The absence of a broadly accepted method or tool for measuring emotion should not preclude emotion research. Nonetheless, it is important to acknowledge that the nature of the empirical research involves relying on proxies and associations, and that empirical research relying nearly exclusively on proxies and associations is not epistemically ideal. There is not an easy fix, at present, for the MFLE research. We think that more research should be done on the connection between the FLE and emotion, and the research that we are discussing is the beginning of that. The best research so far leads us to interesting educated guesses, but guesses nonetheless, at what mediates the relation between emotion and moral judgment.

4. Particularities in Emotions and Their Effects Remain Understudied

It is almost certainly true that there are important connections between emotion and moral judgment when speaking in generalities. However, researchers have attributed subtle causal roles to emotion tokens and types at a fairly fine grain. Consider the following examples

²⁵ One may be concerned that we are setting too high of a standard for empirical research, as alternative explanations are virtually impossible to rule out alternative explanations entirely. It is not an easy task to settle on how strong the correlations and explanatory power need to be to make such inferences. But, in this case, there are already plausible alternative explanations available. Kahane and colleagues (cf. 2012). have already suggested a plausible alternative explanation – that counter-intuitiveness of the dilemma explains the brain activity and emotional arousal. We think we cannot be accused of setting the bar too high for empirical inquiry in this domain when there are plausible alternative explanations readily available.

²⁶ Cf. Greene (2008).

of experimental findings. Seeing a dirty and disorganized room increases conviction to punish.²⁷ Finding a dime in a phone booth caused people to help passersby more frequently.²⁸ And to use our example of the moral foreign-language effect (or MFLE), thinking in L2 makes people more likely to provide a utilitarian response to the footbridge dilemma.

Researchers often fill in the details about the role of specific emotions in these cases. Seeing the dirty room increases disgust, and that's why you form more severe moral judgments. Finding a dime boosts positive affect, which then more readily initiates helping behavior. Using a non-native language "might stunt emotional processing, attenuating considerations of deontological rules, such as the prohibition against killing,"²⁹ and that's why the MFLE occurs.

How reproducible these effects are remains to be seen.³⁰ Our point applies even if the effects are real and stable. Our concern is that in these cases, we are often told which emotion is doing the work, but we are not provided with a more general theory for how emotions interact with moral judgments, or even judgments more broadly construed. With rare exceptions, we are not told how we know it is *this* particular emotion rather than another, and what testing was done to rule out other emotions as possible confounding variables.³¹ Typically, we are not told how individual differences in traits interact with episodic affect, for example.³² More research is needed that recognizes that forming judgments or performing moral behaviors are complicated processes, perhaps complexes of cognitive, motivational, and affective states, with possibly complementary or oppositional reactive relations among them.

This is problematic because it permits a certain degree of freedom of interpretation by the researchers conducting the studies. These explanations – matching particular emotions to the scenarios – are, to a certain extent, speculative and susceptible to over-fitting. After all, we are provided with no evidence that activation of disgust *always* increases the severity of moral judgment, and we doubt such evidence is forthcoming. So then we are left with little conception of how to know when to expect disgust plays this role and when it will not. It would be helpful to have a full-fledged theory that predicts harsher moral judgments across many distinct contexts, or that with a plausible degree of parsimony predicts when we will see the effects of disgust and we will not, one that would bar the possibility of *post hoc* gerrymandering. An important question to answer is how we can tell when disgust, for instance, is the causal agent and when it isn't, prior to finding an effect in a lab experiment. A larger predictive theory can help provide some answers.

²⁷ Schnall, Haidt, Clore et al. (2008).

²⁸ Isen, Levin (1972).

²⁹ Hayakawa, Tannebaum, Costa et al. (2017): 1394.

³⁰ While Batson, Coke, Chard et al (1979) and others have replicated Isen, Levin (1972)'s results, Blevins, Murphy (1974) and Weyant, Clark (1976) did not. Also, see Landy, Goodwin (2015) and Schnall, Haidt, Clore et al. (2015)'s response for some controversy regarding the incidental disgust findings.

³¹ Newer research has been better in this respect – cf. Baron, Gürçay, Luce (2018); Gawronski, Conway, Armstrong et al. (2018); Reynolds, Conway (2018).

³² There are exceptions to all of these generalizations. But they often serve as exceptions that prove the rule. For every study that finds, for instances, that low empathy is correlated with increased utilitarian responding (Choe, Min 2011), another study that shows increased episodes of sympathy are correlated positively with increased utilitarian responding is sure to follow – Baron, Gürçay, Luce (2018).

As we mentioned in Section 2, the authors who originally discovered the foreign-language effect in a non-moral context claim that emotion is tied to Type 1 processing, and Type 1 processing is tied to various heuristics and biases. In the case of loss aversion, one hypothesis is that emotional triggers of fear of loss engage Type 1 processing. Keysar, Hayakawa, and An cite evidence from the neuroscientific literature on decision-making:

There is evidence that myopic loss aversion is driven by an emotional reaction to the prospect of loss, as patients with focal lesions in areas that regulate emotions, compared with nonbrain-damaged participants, are less likely to show this effect, and more likely to accept . . . positive-expected-value bets.³³

When using L2, however, the emotional responses are not triggered, and Type 2 processing is engaged uninterrupted, hence “making a decision in a foreign language could reduce the emotional reaction thereby reducing bias.”³⁴

So, disgust is tied to stronger moral judgments, and perhaps fear influences distinctively deontological judgments. Do fear and disgust activate Type 1 processing? We might guess why these emotions have this causal role. Is it because they are negatively valenced emotions, ones that instruct us to stop what we are doing or avoid the relevant stimulus? Perhaps we should adopt the following hypothesis: negatively valenced emotions trigger Type 1 processing, and neutral or positively valenced emotions do not. (To be clear, we are extrapolating beyond what our interlocutors have claimed in guessing at the hypothesis. Our intention is to think about what it would look like to implement a theory of emotion rather than look at piecemeal cases in the absence of a larger theoretical framework.)

There is at present some research that provides mixed results for this hypothesis. Baron, Gürçay, & Luce (2018) find that while disgust is negatively correlated with utilitarian thinking, we don't have evidence that other emotions (e.g. anger) are negatively correlated with utilitarian judgments.³⁵ In fact, they found a weak positive correlation between anger and utilitarian judgments, which is contrary to what we should expect if utilitarian judgments are the product of Type 2 processing. So, even discussing emotions as positively or negatively valenced might not be fine-grained enough. To further complicate this picture, Gawronski et al. (2018) found no correlation between incidental sadness or anger and moral judgment.³⁶ More research into the distinctive contributions of emotion types to moral judgments is necessary to have confidence in any single study.

³³ Keysar, Hayakawa, An (2012): 665.

³⁴ Ibidem: 667.

³⁵ Baron, Gürçay, Luce (2018) looked for correlations between a range of emotions and traits with utilitarian judgments. They did not find evidence of positive correlations between emotion and deontology or between utilitarian judgments and other emotions.

³⁶ Gawronski, Conway, Armstrong et al. (2018) also found that incidental happiness was correlated with increased utilitarian responding, but that this result was caused by a weaker sensitivity to moral norms. This paper draws on the CNI model, developed in Gawronski, Armstrong, Conway et al. (2017) that attempts to separate the unique contributions of utilitarian reasoning (sensitivity to consequences or outcomes), deontological reasoning (sensitivity to norms), and a general bias toward inaction (Consequences, Norms, Inaction = CNI). If this model is correct, then the picture only becomes more complex as a positive correlation between utilitarian responses and incidental happiness might not be caused by being more responsive to consequences but, instead, by being less responsive to deontological reasons.

Furthermore, the hypothesis that Type 1 processing is activated by negatively valenced emotion, does not be coherent with other areas of psychological research, unrelated to moral judgment. Daniel Kahneman, following Schwarz & Bless,³⁷ writes:

These findings add to the growing evidence that good mood, intuition, creativity, gullibility, and increased reliance on System 1 form a cluster. At the other pole, sadness, vigilance, suspicion, an analytic approach, and increased effort also go together. A happy mood loosens the control of System 2 over performance: when in a good mood, people become more intuitive and more creative but also less vigilant and more prone to logical errors. Here again, as in the mere exposure effect, the connection makes biological sense. A good mood is a signal that things are generally going well, the environment is safe, and it is all right to let one's guard down. A bad mood indicates that things are not going very well, there may be threat, and vigilance is required. Cognitive ease is both a cause and a consequence of a pleasant feeling."³⁸

Kahneman is not intending to weigh in on debates on moral judgment, but if his general theory of the relation between emotion and cognitive processing is correct, we need to re-think the relationship between negatively valenced emotions and Type 1 processing, and hence deontological moral judgments.

Researchers can always make further distinctions. Perhaps not all negatively valenced emotions have equivalent effects on moral judgment. Lerner et al. have developed an "appraisal-tendency" framework which marks differences between negatively valenced emotions and their relationship to judgment and decision-making.³⁹ Anger is tied to a perception of low risk and a conception of events as under one's control. Fear, on the other hand, is tied to a perception of high risk and a conception of events as unpredictable. Ought we predict one or the other tied to deontology over utilitarianism? Does anger trigger Type 1 processing but fear not? Or vice versa? We are lacking a principled reason to conclude in favor of any of these hypotheses. Pfister and Böhm analyze the connection between emotion and judgment according to a different schema, where valence is not particularly important at all.⁴⁰ One or more of these views may turn out to be at least partially correct, and that is the point we are trying to make. There are too many possibilities still in play and some consistency among findings on emotion research and moral judgment is crucial before drawing many firm conclusions.

The evidence used most often to support the idea that moral judgments in L2 involve a reduction of emotion is the research led by Catherine Caldwell-Harris. While there is significant literature on the relationship between bilingualism and emotion, it is not as simple as a reduced emotionality across the board. The relationship between emotion and bilingualism is complex. While some researchers have concluded that all emotional reactions are reduced in a non-native language, this conclusion is

³⁷ Schwartz, Bless (1991).

³⁸ Kahneman (2011): 68-69.

³⁹ Lerner, Li, Valdesolo et al. (2015).

⁴⁰ Pfister, Böhm (2008).

controversial.⁴¹ Caldwell-Harris and colleagues⁴² show that some words and phrases are more emotionally intense in participants' L1 (e.g. childhood reprimands), but that participants still experience emotion in L2 contrary to their verbal assertions that they "felt nothing" in their non-native language.⁴³ Skin conductance responses (SCRs) to taboo words were not significantly higher in their L1, and participants' SCR results were higher for taboo words in their L2 than other emotional words (e.g., pleasant, etc.) in their L1.⁴⁴ While there are potential explanations for this discrepancy between qualitative reports and SCRs, the fact that SCR results were not significantly higher for taboo words in one's L1 seems to indicate that thinking in one's second language is not always less emotional. The relationship between bilingualism and emotion is a complex one, and there is still more to learn.

Returning to the MFLE, the research on emotion and bilingualism as well as the dual-processing research has been used as evidence to support the reduced-emotionality account. Researchers state that Caldwell-Harris's body of research establishes that individuals experience less emotion in their non-native language.⁴⁵ Hence, it makes sense that they would reason differently and make different moral judgments. But, given the complexity of this research, this discussion needs to be more nuanced.⁴⁶ The impact of using a non-native language on emotion depends on the *type of emotion* being studied, along with the method of studying the reaction. And taking a closer look at Kahneman and Lerner's work shows that rather than assuming that emotion is a monolithic force that either is or is not responsible for deontological judgments and the FLE, we need to look at emotion and its effect on moral judgment more carefully and investigate which particular emotions are relevant to moral judgment (or the MFLE) and under what circumstances.

5. Obscurity of Neutral, Non-emotional State

Finally, we address an issue that we believe is further under the radar than those we have discussed so far – the lack of a clear control for emotion research, particularly in the context of research on moral judgment. To start, consider the following elementary points about the reasoning behind experimental methods to determine causal relationships. In simplistic terms, if we are to understand the causal influence of some purported cause *c* on an event *e*, we will also need to understand what happens with *e* when we do not have *c*. We do not mean to be taking a controversial stance on the metaphysics

⁴¹ See De Langhe, Puntoni, Fernandes et al. (2011): 367.

⁴² Harris, Ayçiçeği, Gleason (2003; 2006); Caldwell-Harris, Ayçiçeği-Dinn (2009).

⁴³ Harris, Ayçiçeği, Gleason (2003): 572.

⁴⁴ Ibidem: 572-573.

⁴⁵ See Keysar, Hayakawa, An (2012): 661; Corey, Hayakawa, Foucart et al. (2017): 3; Hayakawa, Tannebaum, Costa et al. (2017): 1388.

⁴⁶ Additionally, there are other potential explanations for these results. In Gawinkowska, Paradowski, Bilewicz (2013), the reduced emotionality explanation for these effects was contrasted with an alternative explanation – that the effect is due to social norms being less vivid in one's L2 rather than reduced emotionality. They only found an effect in cases of "normatively protected expletives" which led them to conclude that "it is social norms and limitations that really motivates bilinguals to swear in L2" (p. 5).

of causation, such as subscribing to a counterfactual account. We are just pointing to a virtual truism regarding causal reasoning. Carrying over to experimental contexts, randomized controlled trials are often viewed as the gold standard for determining the causal effects of some relevant causal stimulus. The control group serves as the experimental demonstration of what happens with *e* in the absence of *c*.

Turn now to emotion research in moral psychology. If *c* = emotion and *e* = the formation of moral judgments, we can ask what happens with *e* in the absence of *c*? What does it mean to say that there is an absence of emotion?

Return to the proposed explanation of the foreign-language effect – the reduced emotionality account. This account states that considering a moral dilemma in one’s non-native language reduces emotion. But, to properly assess the account, there must be some identifiable baseline of emotionality against which we can measure some reduction. What is this baseline level of emotion of comparison? If the reduced emotionality account is correct, we need to know which emotions are reduced, but also by *how much*, and relative to what.

One potential answer is that considering moral dilemmas in one’s non-native language involves a reduction in emotionality relative to the kind of ordinary attitudes the participants were experiencing prior to thinking in L2. To illustrate, call time T_1 the moment before the participant receives the moral dilemma stimulus in his/her non-native language and call time T_2 the moment the participant forms the intuition (this is probably best understood as a duration, but for the sake of uniformity of variables, we’ll treat it as a moment). This potential answer would imply that when the moral dilemma stimulus is considered in L2, the participant will be emotionally stunted at T_2 relative to T_1 .

But the reduced emotionality account cannot be interpreted in this way. It does not seem wise to interpret the reduced emotionality account as committed to the view that considering the moral dilemma in L2 reduces emotion at time T_2 (post-stimulus forming of moral intuition) relative to time T_1 (pre-stimulus). Presumably, participants are not in a particularly elevated state before being presented with the stimulus, and so do not experience a drop off after reading it in L2.

The most natural interpretation of the reduced emotionality account’s claim that “a [non-native language] context reduces the emotional reactivity that certain scenarios may elicit”⁴⁷ is that one would *expect* an emotional reaction as demonstrated in parallel experiments considered in L1, but the elevated emotionality is not present when the dilemma is considered using L2. The typical emotional triggers or cues in the prompt fail to activate emotional processing under the non-native language condition. The difference between individuals’ reaction to the trolley case and the footbridge case is, according to this view, that the idea of pushing an individual onto the tracks (rather than pulling a lever to achieve arguably the same outcome) upsets us. The case “triggers” an emotional reaction that the original trolley case does not. So, rather than reading these cases in L2 reducing one’s current emotional state, it seems that researchers are saying that reasoning in L2 blocks a normal emotional reaction. Hence, *reduced* emotionality is relative to an expected counterfactual baseline of *higher than normal* emotionality found in the L1 condition. Were we just to look at the within subject series of actual events, we

⁴⁷ Vives, Aparici, Costa (2018): 2.

might understand the reduced emotionality account as working via the *absence* of sufficient emotion to activate emotional processing in the actual sequence studied in these experiments. However, virtually all of the experiments conducted on the FLE have been between subject, and individual participants' baseline, pre-stimulus reasoning type or emotion/emotional reactivity status has not been adequately measured.

To summarize, if, when considering a moral dilemma in L2, participants are in the same emotional state at T_2 as at T_1 , this would be taken as evidence in support of the reduced emotionality account. The evidence provided that they are in similar emotional states is their proclivity for utilitarian responding. We think more needs to be done to verify the emotional states, both at T_1 and T_2 , in order to justify this inference. Taken in combination with concerns of diversity of measurement and the use of proxies, the current concern that we are inferring the absence of emotion from the absence of a proxy (deontological response) takes on added weight.

6. Conclusion

The affective revolution has advanced our understanding of cognition, and the empirical study of moral psychology is fruitful and accelerating in this respect. Still, all things considered, we think the confidence in our empirical understanding of the role of emotion in moral psychology has far exceeded what is reasonably licensed. We think that the recognition of the possible distortionary effects of using proxies, more principled attempts to limit degrees of researcher freedom in discerning which emotion is at work under which conditions, and accounting for the obscurity of a control group's emotional state can help advance the foreign-language effect (FLE) research as well as the science of moral psychology more generally.⁴⁸

Acknowledgements

This research has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 805498).

References

- Bago B., De Neys W. (2019), "The Intuitive Greater Good: Testing the Corrective Dual Process Model of Moral Cognition," *Journal of Experimental Psychology: General* 148 (10): 1782–1801.
- Baron J., Gürçay B., Luce M.F. (2018), "Correlations of Trait and State Emotions with Utilitarian Moral Judgements," *Cognition and Emotion* 32 (1): 116–129.
- Batson C.D., Coke J.S., Chard F. et al (1979), "Generality of the 'Glow of Goodwill': Effects of Mood on Helping and Information Acquisition," *Social Psychology Quarterly* 42 (2): 176–179.

⁴⁸ We would like to thank Michał Białek, Kyle Fritz, and John Stigall for providing helpful comments on the manuscript.

- Białek M., De Neys W. (2016), "Conflict Detection During Moral Decision-Making: Evidence for Deontic Reasoners' Utilitarian Sensitivity," *Journal of Cognitive Psychology* 28 (5): 631–639.
- Białek M., De Neys W. (2017), "Dual Processes and Moral Conflict: Evidence for Deontological Reasoners' Intuitive Utilitarian Sensitivity," *Judgment and Decision Making* 12 (2): 148–167.
- Białek 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.
- Blevins G.A., Murphy T. (1974), "Feeling Good and Helping: Further Phonebooth Findings," *Psychological Reports* 34 (1): 326.
- Bond M.H., Lai T.M. (1986), "Embarrassment and Code-Switching Into a Second Language," *Journal of Social Psychology* 126 (2): 179–186.
- Caldwell-Harris C.L., Ayçiçeği-Dinn A. (2009), "Emotion and Lying in a Non-Native Language," *International Journal of Psychophysiology* 71 (3): 193–204.
- Choe S.Y., Min K. (2011), "Who Makes Utilitarian Judgments? The Influences of Emotions on Utilitarian Judgments," *Judgment and Decision Making* 6 (7): 580–592.
- Cipolletti, H., McFarlane, S., Weissglass, C. (2016), "The Moral Foreign-Language Effect," *Philosophical Psychology* 29 (1): 23–40.
- Colbeck K.L., Bowers J.S. (2012), "Blinded by Taboo Words in L1 but not L2," *Emotion* 12 (2): 217–222.
- Corey J.D., Hayakawa S., Foucart A. et al. (2017), "Our Moral Choices are Foreign to Us," *Journal of Experimental Psychology: Learning, Memory, and Cognition* 43 (7): 1109–1128.
- Costa A., Foucart A., Hayakawa S. et al. (2014), "Your Morals Depend on Your Language," *PLoS One* 9 (4): e94842.
- De Langhe B., Puntoni S., Fernandes D. et al. (2011). "The Anchor Contraction Effect in International Marketing Research," *Journal of Marketing Research* 48 (2): 366–380.
- Dewaele J.M. (2010), *Emotions in Multiple Languages*, Palgrave Macmillan, Basingstoke, England.
- Drake B. (2012), "To Avoid Stupid Mistakes," *Bloomberg Report*, URL = <https://www.bloomberg.com/news/articles/2012-04-25/to-avoid-stupid-mistakes-think-in-french> [Accessed 13.02.2020].
- Gawinkowska M., Paradowski M.B., Bilewicz M. (2013), "Second Language as an Exemptor from Sociocultural Norms: Emotion-Related Language Choice Revisited," *PloS One* 8 (12): e81225.
- Gawronski B., Armstrong J., Conway P. et al. (2017), "Consequences, Norms, and Generalized Inaction in Moral Dilemmas: The CNI Model of Moral Decision-Making," *Journal of Personality and Social Psychology* 113 (3): 343–376.
- Gawronski B., Conway P., Armstrong J. et al. (2018), "Effects of Incidental Emotions on Moral Dilemma Judgments: An Analysis Using the CNI Model," *Emotion* 18 (7): 989–1008.
- Geipel J., Hadjichristidis C., Surian L. (2015), "The Foreign Language Effect on Moral Judgment: The Role of Emotions and Norms," *PLoS One* 10 (7): e0131529.
- Geipel J., Hadjichristidis C., Surian, L. (2016), "Foreign Language Affects the Contribution of Intentions and Outcomes to Moral Judgment," *Cognition* 154: 34–39.
- Greene J.D., Sommerville R.B., Nystrom L.E. et al. (2001), "An fMRI Investigation of Emotional Engagement in Moral Judgment," *Science* 293 (5537): 2105–2108.

- Greene J.D., Sommerville R.B., Nystrom L.E. et al. (2008), "The Secret Joke of Kant's Soul" [in:] *Moral Psychology Vol. 3: The Neuroscience of Morality: Emotion, Disease, and Development*, W. Sinnott-Armstrong (ed.), The MIT Press, Cambridge (MA): 35–79.
- Greene J.D., Sommerville R.B., Nystrom L.E. et al. (2013), *Moral Tribes: Emotion, Reason and the Gap Between Us and Them*, Penguin Books, New York.
- Greene R.L. (2012), "Oprima Dos for Better Cognition," *The Economist*, URL = economist.com/johnson/2012/05/08/oprima-dos-for-better-cognition [Accessed 13.02.2020].
- Gross J. (2012), "Reasoning is Sharper in a Foreign Language," *Scientific American*, URL = scientificamerican.com/article/reasoning-is-sharper-in-a-foreign-language/ [Accessed 13.02.2020].
- Haidt J. (2001), "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment," *Psychological Review* 108 (4): 814–834.
- Harris C.L., Ayçiçeği A., Gleason J.B. (2003), "Taboo Words and Reprimands Elicit Greater Autonomic Reactivity in a First than in a Second Language," *Applied Psycholinguistics* 24 (4): 561–579.
- Harris C.L., Ayçiçeği-Dinn A., Gleason J.B. (2006), "When is a First Language More Emotional? Psychophysiological Evidence from Bilingual Speakers", [in:] *Bilingual Minds: Emotional Experience, Expression and Representation*, A. Pavlenko (ed.), Multilingual Matters, Clevedon, England: 257–283.
- Hayakawa S., Tannebaum D., Costa A. et al. (2017), "Thinking More or Feeling Less? Explaining the Foreign-Language Effect on Moral Judgment," *Psychological Science* 28 (10): 1387–1397.
- Horne Z., Powell D. (2016), "How Large Is the Role of Emotion in Judgments of Moral Dilemmas?", *PLoS ONE* 11 (7): e0154780.
- Kahane G., Shackel N. (2008), "Do Abnormal Responses Show Utilitarian Bias?," *Nature* 452: E5.
- Kahane G., Shackel N. (2010), "Methodological Issues in the Neuroscience of Moral Judgment," *Mind & Language* 25 (5): 561–582.
- Kahane G., Wiech K., Shackel N. et al. (2011), "The Neural Basis of Intuitive and Counter-intuitive Moral Judgment," *Social Cognitive and Affective Neuroscience* 7 (4): 393–402.
- Kahneman D. (2013), *Thinking, Fast and Slow*, Farrar, Straus and Giroux, New York.
- Keysar B., Hayakawa S., An S. (2012), "The Foreign-Language Effect: Thinking in a Foreign Tongue Reduces Decision Biases," *Psychological Science* 23 (6): 661–668.
- Landy J.F., Goodwin G.P. (2015). "Does Incidental Disgust Amplify Moral Judgment? A Meta-Analytic Review of Experimental Evidence," *Perspectives on Psychological Science*, 10 (4): 518–536.
- Lerner J.S., Li Y., Valdesolo P. et al. (2015), "Emotion and Decision Making," *Annual Review of Psychology* 66: 799–823.
- Lieberman M. (2017), "Thinking in a Second Language Makes You More Logical," *Travel & Leisure*, URL = travellandleisure.com/travel-tips/languages-affect-our-rationality-and-morality [Accessed 12.05.2020].
- Muda R., Niszczota P., Białek M. et al. (2018), "Reading Dilemmas in a Foreign Language Reduces Both Deontological and Utilitarian Response Tendencies," *Journal of Experimental Psychology. Learning, Memory, and Cognition* 44 (2): 321–326.
- Pfister H., Böhm G. (2008), "The Multiplicity of Emotions: A Framework of Emotional Functions in Decision Making," *Judgment and Decision Making* 3 (1): 5–17.

- Reynolds C.J., Conway P. (2018), "Not Just Bad Actions: Affective Concern for Bad Outcomes Contributes to Moral Condemnation of Harm in Moral Dilemmas," *Emotion* 18 (7): 1009–1023.
- Schnall S., Haidt J., Clore G.L. et al. (2008), "Disgust as Embodied Moral Judgment," *Personality & Social Psychology Bulletin* 34 (8): 1096–109.
- Schnall S., Haidt J., Clore G.L. et al. (2015), "Landy and Goodwin (2015) Confirmed Most of Our Findings Then Drew the Wrong Conclusions," *Perspectives on Psychological Science* 10 (4): 537–538.
- Schwarz N., Bless H. (1991), "Happy and Mindless, but Sad and Smart? The Impact of Affective States on Analytic Reasoning," [in:] *Emotion and Social Judgments*, J.P. Forgas (ed.), Pergamon Press, Elmsford, NY: 55–71.
- Sevinc G., Spreng R.N. (2014), "Contextual and Perceptual Brain Processes Underlying Moral Cognition: A Quantitative Meta-Analysis of Moral Reasoning and Moral Emotions," *PLoS ONE* 9 (2): e87427.
- Valdesolo P., DeSteno D. (2006), "Manipulations of Emotional Context Shape Moral Judgment," *Psychological Science* 17 (6): 476–477.
- Vives M., Aparici M., Costa A. (2018), "The Limits of the Foreign Language Effect on Decision-Making: The Case of Outcome Bias and the Representativeness Heuristic," *PLoS ONE* 13 (9): e0203528.
- Weyant J., Clark R.D. (1976), "Dimes and Helping: The Other Side of the Coin," *Personality and Social Psychology Bulletin* 3 (1): 107–110.