THE UNIVERSITY OF CHICAGO

WHY USING A FOREIGN LANGUAGE CHANGES OUR CHOICES

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

BY

SAYURI HAYAKAWA

CHICAGO, ILLINOIS

AUGUST 2017

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my gratitude to Catherine Caldwell-Harris who first introduced me to research, and whose investigations into emotion and language continue to influence my work today. Howard Nusbaum, who has deepened my understanding of so many things ranging from attentional processes to Aristotle. Greg Norman, whose expertise, good humor, and unyielding patience enriched my graduate experience. Albert Costa, an incredible collaborator and mentor, who has taught me to be skeptical of crazy effects without losing the courage to seek them out. David Tannenbaum, Elizabeth Necka, and Shannon Heald, who patiently shared their expertise with me. I give my thanks to the William Rainey Harper Dissertation Year Fellowship, the John Templeton Foundation, and the NSF for financial support. I would also like to thank María Cecilia Lozada for her enormous assistance in accessing participants as well as my many wonderful RAs, collaborators and genius undergrads, including Jaewon, Sophie, Beatrix, Patricia, Sylvia, Angel, Joey, and my girl Joanna, who were indispensable in helping me run my studies.

James and Naveen, my big brothers of the psych department, who taught me so many things both in and outside the lab. Sheila, Leigh and Becky, my partners in crime, whose laughter could always be heard in the hallway, making Beecher Hall one of the happiest places I know. Linda Ginzel, the teacher and person that I wish I could be. Whose support made it so even the biggest problems never felt unmanageable, and even the smallest accomplishment never went uncelebrated. I cannot thank her enough for all that she has done for me.

To my advisor, Boaz Keysar, who has taught me everything I know. He taught me to avoid jargon and speak to people as if they were people. He inspired me to seek out research that is not only rigorous, but above all, interesting. He has taught me that sometimes the best use of the day is to talk myself out of a project, and to celebrate a disaster well averted. Who, by example, has taught me to be a kinder and more understanding person. I will carry these lessons with me throughout my life and hope that someday, I can do for someone what he has done for me. From the bottom of my heart, thank you.

Alex, who inspires me to keep learning new things every day and who supplies me with YouTube videos of mini pigs to counteract my mini panic attacks. I could not have done this without you. Lastly, my parents, Linda and Hajime, and my brother and sister-in-law, Ken and Yoshiko. My amazing family and biggest cheerleaders, who freak out with every publication and never for a second let me doubt their unconditional support. I am forever indebted for all of the sacrifices you have made to get me where I am, and I just hope that I can make you proud.

ABSTRACT

Imagine you are faced with a moral dilemma. Would you expect your actions to differ depending on whether you are using a native or foreign language? As unintuitive as it may seem, using a foreign language changes our choices. While there is increasing evidence of such foreign language effects in domains ranging from moral judgment to risk taking, little is known about the underlying processes. Utilizing both behavioral and physiological measures, the studies presented here explore potential explanations for why using a foreign language affects the decisions we make. The results from Studies 1-6 suggest that those using a foreign language make more utilitarian moral choices, and that this is due to a reduction in emotional processing rather than an increase in analytical processing. Despite the apparent reduction in emotional processing, Studies 7 and 8 reveal no effects of language on either heart rate or heart rate variability when participants were asked to make risky and consumer choices. These results raise the possibility that foreign language use may dampen affective or heuristic processing without necessarily affecting emotional arousal. Contrary to past results, Study 7 also revealed no effect of language on risk taking, suggesting that foreign language effects within this domain may not be as robust as for moral judgments. Study 8 demonstrated that using a foreign language increases both buying and selling prices during a consumer decision task, potentially due to a reduction in the hedonic value of money. Lastly, Studies 9-12 explored the role of mental imagery for moral and risky judgments and find that using a foreign language results in less vivid

visualizations, which partially explain behavioral effects. Taken together, the findings of this project suggest that using a foreign language changes our choices due to a reduction in affective processing, which, in part, results from less vivid mental imagery when using a non-native tongue.

INTRODUCTION

"Don't you see that the whole aim of Newspeak is to narrow the range of thought? In the end we shall make thoughtcrime literally impossible, because there will be no words in which to express it." – George Orwell, '1984'

In Orwell's dystopian 1984, the authoritarian government reinvents language to restrict freedom of thought. Indeed, the word "free" itself could no longer be applied to concepts such as "political freedom" or "intellectual freedom" and would only make sense in the context of statements such as "the dog is free from lice" (Orwell, 1949). The goal of this endeavor was not only to prevent individuals from speaking of controversial topics, but to make the very ideas inconceivable. Indeed, there is significant evidence in support of this notion that language can mold the boundaries of thought and shape the way we experience the world. For instance, the languages we speak can affect how we discriminate between colors (Winawer et al., 2007), perceive time (Boroditsky, 2001; Casasanto, 2008), experience music (Dolscheid et al., 2013), and categorize objects (Lucy, 2004). In all of these cases, researchers have examined whether particular languages (e.g. English versus Russian) result in different patterns of thought. However, there is a relatively unexplored aspect of language that can affect how we think, feel, and even behave. That is, whether the language is processed in a native or foreign tongue.

In this increasingly globalized world, millions of individuals utilize a foreign language every day. Foreign-born doctors make important medical decisions, businesses negotiate international deals, and representatives from around the world attempt to come to agreements on issues ranging from human rights to climate change. While factors such as personal interests, ideologies, and cultural differences easily spring to mind as critical to how these events will

transpire, a less obvious component is the nativeness of the language in use. There is a growing body of research demonstrating that using a foreign language changes our choices, and yet relatively little work has been done to determine the processes driving such effects. Utilizing both behavioral and physiological measures, a series of 12 experiments explore potential mechanisms underlying such effects in domains ranging from risk taking to moral judgment. These include investigations into whether behavior changes because of reduced emotional processing, increased analytical processing, and reduced mental imagery when using a foreign tongue. In this way, the goal of the present research is to help shed some light on how and why using a foreign language changes the choices that we make.

Keysar, Hayakawa & An (2012) first documented this "foreign language effect" in the domain of risk taking by demonstrating that people randomly assigned to gamble in a foreign tongue were less risk averse relative to those assigned to use their native tongue. In one experiment, the authors presented participants with a series of bets with positive expected value and found that those using their native language only accepted an average of 54% of the bets, while this number went up to 71% when using an understood, but foreign language. Costa and colleagues (2014a) found a similar result when presenting participants with the Holt-Laury test (Holt & Laury, 2002) which involves making ten choices between pairs of lotteries. The researchers found that those using a foreign language were more willing to take an advantageous risk in cases where one choice did not clearly dominate the other. The apparent increase in risk taking when using a foreign language may, in part, be explained by differences in risk perception. Hadjichristidis, Geipel, and Savadori (2015) presented participants with various potential hazards such as "nanotechnology" and asked them to make judgments of perceived

benefits and risks. Those using a foreign language perceived greater benefits and less risk relative to those using their native tongue.

While both Keysar et al. (2012) and Costa et al. (2014a) found a foreign language increase in risk taking, Gao and colleagues (2015) actually observed a foreign language decrease. One potential explanation may be that Gao et al. included gambles with negative expected value whilst the previous experiments included only those with positive expected value. In addition to finding that using a foreign language affects risk differently depending on the context, Gao et al. demonstrated that using a foreign language reduced the so-called "hot hand" fallacy. Participants were asked to decide whether or not to accept a series of independent 50/50 gambles and were given linguistic feedback (e.g. "Excellent!", "Sorry!") in either their native or foreign language. The authors found that those using their native language were significantly more likely to accept a gamble after receiving positive feedback in the previous round, despite the gambles being independent, whereas those using their foreign language were not affected by feedback. This suggests that individuals may be less swayed by irrelevant information when making risky choices in a foreign language.

Indeed, there is consistent evidence that using a foreign language does lead to decisions that are less biased by irrelevant information. Keysar et al. (2012) presented participants with the "Asian Disease" problem (Tversky & Kahneman, 1981) in either their native or foreign language. This problem illustrates the well-known “framing effect” which describes how individuals tend to be risk averse when a decision is framed in terms of the potential gains (e.g. 1/3 of the people will live) while they become risk seeking when the same exact problem is simply described in terms of potential losses (e.g. 2/3 of the people will die). The authors found that this very robust effect was significantly reduced or even eliminated when participants made

the choice in a foreign language. Similar results were obtained by Winskel et al. (2016) as well as Costa et al. (2014a). Nakamura (2016) varied the procedure by employing a fully within-subject design with all participants seeing both frames in both languages. When utilizing this design, Nakamura did not replicate the foreign language effect for the Asian Disease problem, but did replicate it for a financial variant of the problem. This suggests that the effect is less robust, though still present when participants are exposed to both languages and both frames.

Using a foreign language has been shown to affect decision making in domains outside of risky choice. For example, Urbig and colleagues (2016) found that those using a foreign tongue were more likely to freeride when presented with a public-goods game in which participants could contribute to a communal pot which was then divided among all players (e.g. Rand, Greene & Nowak, 2012). Díaz-Lago & Matute (2014) found that people using a foreign language were less likely to incorrectly infer a causal relationship between two unrelated events. The most dramatic and robust foreign language effects, however, have been found in the moral domain. Consider the following problem:

> *You are standing on a footbridge overlooking a train track. A small on-coming train is about to kill five people and the only way to stop it is to push a heavy man off the footbridge in front of the train. This will kill him, but save the five people. Would you push the man?*

When participants are presented with this "Footbridge" problem (Foot, 1978; Thompson, 1985), the vast majority of participants refuse to take this emotionally aversive, but utilitarian action of sacrificing one life to save five (Greene et al., 2008; Bartels, 2008). While it may seem surprising that such a serious life-or-death decision would be affected by something as unrelated as the language being spoken, there is now substantial evidence that language indeed does matter. Costa et al. (2014b) first discovered that when participants were randomly assigned to answer

this Footbridge dilemma in either their native or foreign language, those using their foreign tongue were sometimes as much as twice as likely to opt for the utilitarian action of sacrificing one life to save five. This effect has been replicated with a number of different language pairs in different labs around the world (e.g. Geipel, Hadjichristidis, & Surian, 2015a; Nakamura, 2015; Nakamura, 2016; Cipolletti, McFarlane & Weissglass, 2016; Corey et al., 2017; Im Shin & Kim, 2017).

In addition to the Footbridge dilemma, Costa and colleagues presented participants with the relatively less emotional "Switch" variant of the problem. In this case, the only way to save the five people is to flip a switch that will divert the train to a different track where one person will be killed. While the final outcomes are identical, this less aversive means of sacrificing the one person usually leads to significantly more endorsement of the utilitarian action (e.g. Greene et al., 2001). In this case, no foreign language effect was found with the vast majority of both language groups opting for the utilitarian choice. This pattern is consistent with the authors' proposed explanation for the effect, which is that those using a foreign language are less emotional than those using their native tongue, thereby leading to more utilitarian behavior specifically when the dilemma is highly emotionally aversive. Similarly, Geipel, Hadjichristidis and Surian (2015b) found that using a foreign language leads to more lenient judgments of taboo and social norm violations, such as eating a dead dog and consensual incest, but only for more emotional situations.

Foreign language effects on moral judgment and decision-making have also been observed when utilizing scenarios that involve what may be considered more "real world" contexts. For instance, Geipel et al. (2015b) found that those using a foreign language were more lenient when judging actions such as selling someone a defective car or cutting in line when in a

hurry. Pan and Patel (2016) asked accounting students to recommend what financial information to disclose in a report to examine the effect of language on ethical behavior. The authors found that those using a foreign language were more likely to endorse selectively disclosing information for personal gain and suggest that this may stem from a detachment from norms and deontological rules. These more ecologically valid scenarios suggest that using a foreign language may have real-world consequences for the millions of individuals who regularly utilize non-native languages. If this is the case, it is critical to gain a better understanding of why language changes our choices.

While there have been some explorations into potential mechanisms, there has yet to be a definitive explanation for why our choices change when we use a non-native tongue (see Hayakawa et al., 2016; Hadjichristidis, Geipel & Surian, 2016; and Costa, Vives & Corey, 2017 for discussions of potential mechanisms). Before expanding on some likely mechanisms, I will briefly review a number of explanations that have been explored, but are less likely to be true.

The first is that the various foreign language effects are a result of simple lack of comprehension and therefore random responding when using a foreign language. While this explanation cannot be definitively ruled out for each and every foreign language effect, there are reasons to think this is not the case. Considering the most commonly documented Footbridge foreign language effect, random responding should result in regression to the mean for both Footbridge and Switch dilemmas. However, studies have consistently demonstrated that foreign language effects emerge exclusively for the Footbridge and not for the Switch, suggesting systematic effects of language (e.g. Costa et al., 2014b; Geipel et al., 2015a; Cipolletti et al., 2016; Corey et al., 2017; but see Nakamura, 2015 who did observe what appeared to be regression towards the mean).

Another potential explanation is that these effects arise not because of using a native versus foreign language, but rather specific languages (e.g. an "English effect"). Indeed, there is research suggesting that bilingual biculturals may display different behaviors depending on which language they are speaking, not because one is foreign, but rather because different languages activate different cultural scripts, norms and values (e.g. Benet-Martínez et al., 2002; Ramírez-Esparza et al., 2006). While such a mechanism may exist orthogonally to the foreign language effect, it is unlikely to explain it as multiple experiments have demonstrated similar effects even when languages are completely crossed. For instance, Costa et al. (2014b) as well as Cipolletti et al. (2016) found that using a foreign language increases utilitarianism both when the native language is Spanish and the foreign language is English as well as with the opposite pairing.

Lastly, another potential explanation concerns the inferences made regarding the social membership of the individuals involved in the various dilemmas. Participants may infer that the scenarios involve in-group members when using a native language, but outgroup members when using a foreign language, resulting in relatively greater psychological distance in the case of the latter. To assess this possibility, Corey et al. (2017) presented the Footbridge to native Spanish participants in either Spanish or English and made the membership of the individuals explicitly either American or Spanish. The authors found the usual pattern such that foreign language users were more utilitarian regardless of membership, suggesting that different inferences alone cannot account for the effect. Similarly, Geipel et al. (2015b) obtained a foreign language effect even after clearly stating that the scenarios took place in the participants' native country and involved co-nationals.

Having considered several unlikely explanations, we turn to the leading candidate, which is that using a foreign language changes our choices because it is less emotional relative to a native tongue. According to dual-process models of decision-making, our choices are guided by two different modes of thought. One is intuitive, affective and heuristic (System-1) while the other is analytic, deliberative and relies more heavily on mental resources (System-2; e.g. Kahneman, 2003; Sloman, 1996; Stanovich & West, 2000). Many decision-making biases are thought to be driven by emotional, System-1 processes. For instance, De Martino and colleagues (2006) found that framing effects are associated with activation in emotion-related areas of the brain such as the amygdala. Sanfey and colleagues (2003) similarly found that unfair offers during an Ultimatum game elicited activity in the emotion-related anterior insula. Within the moral domain, Greene et al. (2001) found increased activity in emotion-related regions (e.g. medial frontal gyrus) for moral-personal dilemmas such as the Footbridge relative to less personal dilemmas. This is consistent with findings that individuals with damage to emotion-related areas such as the ventromedial prefrontal cortex are abnormally utilitarian when presented with such problems (Koenigs et al., 2007). As such, one explanation for the various foreign language effects, such as the reduced impact of framing and increased utilitarianism, is a reduction in emotional processing.

Indeed, there is ample evidence suggesting that using a foreign language is less emotional than the native tongue (e.g. Pavlenko, 2005). For example, people have stronger physiological responses to taboo words and reprimands in their native language relative to a foreign language (Harris, Ayçiçegi, & Gleason, 2003). They also rate words and phrases as more emotionally impactful when using their native tongue (e.g., Dewaele, 2004, 2008; Puntoni, de Langhe, & van Osselaer, 2009). Reading emotional passages from Harry Potter elicited weaker and less

differentiated brain activity in emotion-related regions when using a foreign tongue relative to a native language (Hsu, Jacobs & Conrad, 2015). Many more laboratory experiments have demonstrated this foreign language reduction in emotion (see Caldwell-Harris, 2015 and Pavlenko, 2012 for reviews). The question, then, is whether this can explain the foreign language effects on decision-making.

Support for the "reduced emotion" explanation is mixed. As noted previously, researchers have consistently found foreign language effects for the more emotionally aversive Footbridge dilemma while no effects of language were found for the less emotional Switch dilemma (e.g. Costa et al., 2014b; Geipel et al., 2015a; Corey et al., 2017). Additionally, as noted, Geipel et al. (2015b) found that using a foreign language reduces the harshness of moral judgments specifically for more emotional dilemmas. The authors found that in these cases, using a foreign language also reduced ratings of emotional distress, which then mediated the relationship between language and moral judgment. Turning to the domain of risk, Hadjichristidis et al.'s (2015) finding that using a foreign language reduces perceptions of risk and increases perceptions of benefits was mediated by ratings of how positive versus negative participants felt when judging potential hazards. The authors found that participants using a foreign language felt significantly more positively about hazards such as "nanotechnology" and significantly less negatively, resulting in different perceptions of risk and benefit. This asymmetry in the reduction of positive versus negative emotions has been observed before by Wu & Thierry (2012) who found that using a foreign language selectively decreases negative emotions. Also in the domain of risk, Gao et al. (2015) found differential EEG activity while participants made risky choices in their native versus foreign tongue. The Feedback-Related Negativity (FRN), an event-related potential known to index reward sensitivity (e.g. Miltner, Braun & Coles, 1997), was amplified

when participants made choices in their native language as compared to their foreign tongue. Additionally, while the amplitude of the FRN significantly differed following winning versus losing trials in the native language, no such sensitivity was detected in the foreign language.

Lastly, Klesse, Levav, & Goukens (2015) provide some intriguing support for the reduced emotion explanation in the relatively uninvestigated domain of self-control. In a series of experiments, the authors first found that individuals consistently made more indulgent food choices when the choice was expressed orally rather than via button press. The authors predicted this effect based on the theory that oral responses activate emotional processes more strongly than manual responses. Critically, participants making oral choices in a foreign language were no more indulgent than those who made a manual decision, suggesting that the foreign language may not activate emotional responses as much as a native tongue.

All of these results support the idea that previously observed foreign language effects on decisions and judgments may result from a reduction in emotional processing. However, there is some evidence that speaks against this notion. For instance, Geipel et al. (2015a) found that while participants rated both the Footbridge and the Switch dilemmas to be less emotional in the foreign language, these ratings did not mediate participants' choices. A similar result was obtained by Chan et al. (2016).

Further evidence against the "reduced emotion" hypothesis may come from Duñabeitia & Costa (2015) who utilized pupillary dilation as a measure of arousal when examining native and foreign language users telling both veridical and false statements. As expected, greater arousal was detected for lies over truthful statements, but this was not qualified by language, as might be expected if using a foreign language results in less emotion. Indeed, the authors found a separate

effect of language such that using a foreign language elicited *greater* arousal independent of veracity. The authors note, however, that the independence of these two stressors suggest that the stress of lying did not "snow ball" into the stress of speaking a foreign language as one might expect. As such, the lack of interaction may indicate a potential compensatory mechanism when using a foreign language, such as reduced emotion, that prevents the added stress of lying from compounding on the anxiety produced by using a less familiar language. Lazar, Stern & Cohen (2014) offer a similar interpretation for not finding a foreign language reduction in physiological measures of arousal (heart rate variability and skin conductance) and an unexpected increase in prefrontal cortex activity using EEG. The authors suggest that the lack of effect for the physiological measures may be due to heightened anxiety, which could be counteracting the effect of emotional distance when using a foreign language.

While heightened anxiety may mask a foreign language reduction in emotion, other factors may eliminate the emotional blunting altogether. One major factor that is likely to affect both anxiety and emotionality is proficiency in the foreign language. Indeed, this pattern appears to be reflected in the foreign language effects on decision-making. Costa et al. (2014b) found that while both high and low proficiency participants were more utilitarian than those using their native tongue when responding to the Footbridge dilemma, those who were below average in foreign language proficiency were significantly more utilitarian than those who were above average. A similar negative relationship between proficiency and utilitarianism was found by Geipel and colleagues when utilizing the Footbridge dilemma (2015a). Indeed, Čavar & Tytus (2017) did not find a language effect when the participants were highly proficient in the foreign language. Additionally, when Geipel et al. (2015b) presented participants with taboo and social norm violations, they found that not only were those using a foreign language more lenient in

their moral judgments, but also that participants became increasingly lenient with decreased proficiency. All of these results are consistent with the suggestion that increased proficiency in the foreign language may lead to behavior that is more aligned with those using their native tongue (but see Gao et al., 2015 who found no effect of proficiency on risk taking).

While a reduction in emotional System-1 processing could explain many of the foreign language effects on decision-making, an alternative explanation is that using a foreign language increases deliberative System-2 processing. Such an increase could be expected due to an increase in processing disfluency when using a foreign language. Past research has demonstrated that inducing cognitive disfluency can push people to a more deliberative mode of processing as it signals a need to slow down and pay attention (e.g. Alter et al., 2007). Alter and colleagues presented participants with the Cognitive Reflection Test (CRT; Frederick, 2005) in either clear or degraded font to manipulate fluency. The CRT is made up of questions for which the intuitive response leads to an incorrect solution. For example, consider the following problem:

> *A baseball-bat and a baseball-ball cost 1.10 Euros in total. The bat costs one Euro more than the ball. How much does the ball cost?*

The intuitive answer is 0.10 Euros, when the correct answer is in fact 0.05. The extent to which individuals are able to override this intuitive response is believed to be a measure of reflectiveness, and as predicted, those with disfluent font performed better. The authors found that inducing disfluency could also affect judgments such as by reducing the representativeness heuristic (Kahneman & Tversky, 1973) and improving performance on a syllogistic reasoning task. Similarly, Hernandez and Preston (2012) found that participants demonstrated less confirmation bias, leading to less extreme attitudes when reading arguments in a disfluent format. To the extent that using a foreign language is less fluent than a native tongue, we might

expect a similar increase in deliberative processing. Indeed, Favreau & Segalowitz (1983) provide evidence that foreign languages are processed less automatically than a native tongue, especially at lower levels of proficiency.

While it seems intuitive that foreign languages would be processed less fluently than a native tongue, there is yet little direct evidence to support the theory that this leads to an increase in deliberative thinking. One potential indication of such a mechanism, however, was noted by Frey & Gamond (2015) when evaluating Gao et al.'s (2015) research demonstrating a foreign language reduction in the "hot hand" fallacy. As previously noted, Gao and colleagues found a foreign language decrease in the feedback-related negativity when making risky choices, which the authors took as an indication of decreased emotional processing. However, in addition to the effect of language on the FRN, the authors also found a foreign language increase in the P300, which is positively associated with increased resource allocation (e.g. Wu and Zhou, 2009) and was positively correlated with reaction times in Gao's study. Additionally, differences between P300 amplitudes when using the native versus foreign language in this within-subject design were significantly and positively correlated with native vs. foreign differences in risk taking. As such, Frey and Gamond suggest that people using a foreign language may demonstrate less of a "hot hand" bias, not because they are less emotional, but rather because the added effort of using the foreign language may slow them down and move them to a more deliberative mode of thinking.

Oganian, Korn and Heekeren (2016) also questioned the "reduced emotion" account concerning the framing effect and suggest that the language effect may result from increased cognitive control. However, rather than increased disfluency, the authors suggest that this boost in cognitive control may result from the engagement of executive functions necessary for

switching between languages. Previous research suggests that bilinguals may benefit from certain cognitive advantages as a result of having to mentally juggle multiple languages (e.g. Bialystok et al., 2004). Oganian and colleagues thus hypothesized that people assigned to utilize a foreign language in an experiment are more likely to have experienced a language switch from their native language environment, leading to an acute increase in cognitive control. Indeed, the experimenters found that individuals who experienced any language switch, either from the native to foreign or foreign to native, had a reduced framing effect whereas they did not replicate the overall foreign language reduction (but see Corey et al., 2017, Experiment 2a where there was an effect of language, but not switching for the footbridge dilemma). Thus while the proposed routes to increased cognitive control may differ across studies, there has been some evidence suggesting that using, or at least switching to, a foreign language may promote deliberative thinking.

Despite some evidence that may be consistent with a foreign language increase in deliberation, the evidence against such a mechanism appears to be more robust. In order to examine whether using a foreign language leads to more reflective thinking, Costa et al. (2014a) randomly assigned participants to complete the previously described Cognitive Reflection Test (Frederick, 2005) in either their native or foreign language. Contrary to the *increased deliberation* hypothesis, Costa and his colleagues found that the two language groups performed comparably. In accordance with Costa et al., Turula (2016) found that those using a foreign language did not score higher on the CRT. This result seems to speak against the theory that using a foreign language increases reflectiveness. Indeed, as the authors note, it would perhaps be more intuitive to expect a foreign language reduction in CRT scores as both foreign language use and mathematical processes rely on working memory resources (Robinson, 2001; Ashcraft &

Krause, 2007). Geipel, Hadjichristidis, and Surian (2015b) found such a foreign language reduction even when participants were presented with a non-mathematical task. In this case, participants were randomly assigned to either the native or foreign language and were presented with the Moses Illusion task (Song & Schwarz, 2008). For this task, participants were asked: "How many animals of each kind did Moses take on to the ark?" with the correct answer being "can't say" as the biblical story involved Noah, not Moses. The authors found that those using the foreign language were less likely to sidestep the intuitive answer of "two" relative to the native language, speaking against the idea that using a foreign language increases reflectiveness.

Given the mixed results from previous studies, more research is needed to help elucidate the underlying mechanisms driving the various foreign language effects. The *reduced emotion* and *increased deliberation* theories may be related, such that the reduction of one implies an increase in the other. However, a recent meta-analysis of over 80 studies found no reliable correlation between effects driven by intuitive and analytical processes, suggesting that System-1 and System-2 processing at the individual-level are not opposing ends of a single continuum (Wang et al., 2015). As such, it is important to consider how using a foreign language may act on each of these systems independently, as well as uncover their relative contributions to the foreign language effect on decision-making. Doing so will paint a clearer picture of how millions of bilinguals are affected by their languages, and hopefully, provide a greater understanding into the complicated relationship between language, thought, and behavior.

STUDIES 1-6:

THE FOREIGN LANGUAGE EFFECT ON MORAL JUDGMENT

Would you kill one person to save five? Would your answer depend on whether you are using your native or foreign language? A growing body of research suggests that it does. As noted in the introduction, it has been found that using a foreign language increases the willingness to sacrifice one life for the greater good (Cipolletti, McFarlane, & Weissglass, 2015; Costa et al., 2014b; Geipel, Hadjichristidis, & Surian, 2015a). In one study, Costa & colleagues (2014b) found that while only 13% of people were willing to make the sacrifice in their native tongue, this number rose to 36% when using a foreign language. While the effect appears to be robust across languages and contexts, the underlying process is not yet well understood.

Using a foreign language may act on a number of different processes involved in moral decision-making. According to dual process accounts of moral judgment, our choices are driven by at least two processes: the relatively automatic and emotional System-1 and the more deliberative and rational System-2. There is evidence to suggest that deontological, rule-based beliefs such as "killing is wrong" are supported by the more emotional System-1, while utilitarian ideals such as concerns for the greater good are served by the more deliberative System-2 (Green et al., 2001; Green et al., 2008; Koenigs, et al., 2007). Using a foreign language could affect choice by acting on either one or both of these systems.

If using a foreign language is less emotional, then we might expect to see a decrease in deontological concerns. That is, when using a foreign language, people would be more willing to take an action such as sacrificing one life because it is less emotionally aversive, or because rules and norms are less salient (see Geipel, Hadjichristidis, & Surian, 2015b for discussion of norms

in a foreign language). A second possibility is that using a foreign language affects choice by increasing reliance on System-2 processing. If this were the case, then we might expect to see an increase in utilitarian concerns. That is, people will be more willing to sacrifice the man specifically because they are concerned with the greater good. In this case, they need not feel less averse to pushing the man, but rather are more willing to overcome this aversion due to the greater weight placed on serving the greater good.

The *reduced deontology* and the *increased utilitarianism* accounts need not be mutually exclusive. However, traditional analyses of moral decisions do not allow us to distinguish between the two (Conway & Gawronski, 2013). Indeed, different research teams have interpreted similar findings as evidence of two different psychological mechanisms. Costa et al. (2014b) argue that the foreign language effect "stems from the reduced emotional response elicited by the foreign language, consequently reducing the impact of intuitive emotional concerns" (also see Geipel et al., 2015a, for a similar interpretation). On the other hand, Cipolletti et al. (2015) explained the effect emerged because "thinking in one's non-native language activates systematic processing characteristic of Type 2 processing". The goal of this investigation is to experimentally separate these processes in order to understand the role of foreign language use in moral judgment. To do so, we conducted six experiments utilizing a process dissociation technique that disentangles utilitarian and deontological judgment (Conway & Gawronski, 2013; Jacoby, 1991).

Methods

Our experiments varied in terms of language populations and experimental stimuli, but all shared the same basic procedure. For purposes of efficiency, we first describe the basic experimental procedure followed by a discussion of differences between experiments.

Participants and Procedure

All participants were bilingual, with most acquiring their foreign language in a classroom setting. None of our participants grew up speaking their foreign language at home. Participants were randomly assigned to complete the study in either their native or foreign language. All experiments but one (Experiment 3) were conducted online and took place entirely in the assigned language. All materials were translated and back-translated from English to ensure comparability (Brislin, 1970).

Upon completing an initial language background screening, participants were only allowed to proceed to the study if they reported (i) being a native speaker of the target native language, (ii) being a foreign speaker of the target foreign language, and (iii) did not grow up speaking the target foreign language at home. Eligible participants went through a second phase of screening by completing a proficiency quiz that involved reading a paragraph in the assigned language and answering a multiple-choice question about what they had just read. Only participants who answered the question correctly were allowed to participate in the experiment.

After the screening procedure, participants were presented with 20 moral dilemmas (see "Process Dissociation Task" below). Next, participants translated one moral dilemma from the designated language of the experiment to the other language. This was done to ensure attention and comprehension of the target stimulus materials. The dilemma was randomly selected in advance from the stimuli set and was the same for all participants. Participants who failed to translate any part or who wrote gibberish were excluded from the final analysis. All exclusions are listed in Table 1.

Study	Comprehension	Incomplete	Foreign Dominant	Perfect U
1	15	1	0	0
2	4	5	0	0
3	3	0	1	1
4	4	0	3	3
5	2	10	0	0
6	0	15	0	3

Table 1: Participant exclusions. Number of participants excluded for lack of comprehension, incomplete surveys, having a dominant foreign language, or having a perfect utilitarian score making it mathematically impossible to calculate the deontology score, as will be discussed below.

Lastly, participants completed a set of demographic questions and rated their proficiency on 7 point scales for speaking, listening, reading, and writing in their native and foreign languages. Summary statistics for demographic and language background information can be found in Table 2.

Study	Native	Foreign	Female	Age	AOA	Months Abroad	Proficiency -Native	Proficiency -Foreign
1	German	English	49%	38	11	5	6.77	4.97
2	English	Spanish	72%	21	14	4	6.93	5.29
3	Spanish	English	61%	21	12	2	6.57	5.40
4	German	English	42%	35	11	2	6.16	4.84
5	German	English	46%	32	11	4	6.81	5.19
6	English	German	50%	24	16	14	6.95	5.17

Table 2: Demographic information. "AOA" is the age of foreign language acquisition and "Months Abroad" refers to the number of months spent in a country where the target foreign language is the dominant language.

Process Dissociation Task. Participants responded to 20 moral dilemmas in a random order. For each scenario, participants were asked whether a given action was appropriate by

selecting “Yes,” “No,” or “I don’t understand.” We excluded trials where participants selected the “I don’t understand” option (comprising 0.26–1.54% of trials across studies) and performed analyses on the remaining proportion of “Yes” and “No” responses.

We used a set of moral dilemmas designed to provide independent measures of deontological and utilitarian responding for each participant (Conway & Gawronski, 2013). The key feature of this technique is the presentation of 10 “incongruent” and 10 “congruent” moral dilemmas. Traditional moral dilemmas, such as sacrificing one person to save five people, are incongruent in the sense that deontological and utilitarian concerns conflict — deontological concerns prohibit killing a person to save five whereas utilitarian concerns demand it. Congruent dilemmas are structurally identical to incongruent dilemmas except that deontological and utilitarian proscriptions are in agreement. For example, if given the choice of sacrificing one life in order to prevent five people from being mildly injured, neither deontological nor utilitarian concerns would endorse sacrificing the one person.

Comparing response rates for congruent and incongruent dilemmas allows us to recover separate measures of deontological and utilitarian considerations. To do this, we followed the method detailed in Conway and Gawronski (2013). First, we calculated a utilitarianism parameter (*U*) for each participant by taking the difference in the proportion of “unacceptable” responses for congruent trials compared to incongruent trials:

$$U = p(\text{unacceptable}|\text{congruent}) - p(\text{unacceptable}|\text{incongruent}). \quad (1)$$

Thus, participants scoring high on utilitarianism would find harmful actions unacceptable when they fail to maximize net welfare (i.e., congruent trials), but acceptable when they maximize net welfare (i.e., incongruent trials). Those scoring low on *U* would judge harmful actions as comparably acceptable regardless of whether it maximizes net welfare. Scores could range from

–1 to 1 with the mass of the distribution falling between 0 and 1 (negative *U* scores are possible but rare, as they imply that participants would find it acceptable to kill an innocent person to save five from mild harm, but not acceptable to kill an innocent person to save five lives).[1,2]

To arrive at a separate measure of deontological considerations (*D*), it is useful to first consider the proportion of instances in which utilitarianism does not drive responses (1 – *U*). This figure will include deontologically-driven judgments, plus any other response tendencies that lead participants to find an action acceptable in congruent trials. To isolate *D*, we take the proportion of instances where deontology alone is driving response rates (i.e., incongruent dilemmas):

$$D = p(\text{unacceptable}|\text{incongruent})/(1 - U) \quad (2)$$

Thus, *D* can be thought of as what is left over after partialling out both non–utilitarian *and* non–deontological response tendencies. Scores on *D* range from 0 to 1, with higher scores indicating greater deontological responding.

Experimental Permutations

Our six experiments differed along three dimensions: (i) the type of bilingual population used, (ii) how we elicited responses from participants, and (iii) the set of moral dilemmas to which participants responded. We discuss each permutation below; Table 3 provides an overview.

Bilingual Populations. Our experiments differed in terms of participants' native and foreign languages. Participants' native language was either German (Experiments 1, 4, and 5),

[1] Across our six experiments, 40 participants had negative U scores, comprising approximately 6% of the sample. These participants were included in the analysis.

[2] Across our six experiments we had 7 participants who had a *U* score of 1. Because we cannot calculate a corresponding *D* score for these participants (as this would require dividing by 0), we drop these participants from the analysis as recommend by Friesdorf, Conway, and Gawronski (2015).

English (Experiments 2 and 6), or Spanish (Experiment 3). Participants' foreign language was either German (Experiment 6), English (Experiments 1, 3, 4, and 5), or Spanish (Experiment 2). Collectively, our experiments allow us to examine whether effect sizes vary according to specific native or foreign languages. In some cases we can also compare experiments where native and foreign languages are fully crossed (e.g., comparing the effect of native German speakers responding in English to native English speakers responding in German as in Experiments 5 and 6). Doing so allows us to cleanly disentangle whether our results are driven by using a foreign language rather than using a specific language.

Elicitation Format. Our experiments differed in how participants provided their responses. In Studies 1 and 2, participants were asked questions of the form *"Is it appropriate to push the man off of the bridge?"* similar to the elicitation format used by Conway and Gawronski (2013). For Studies 3 and 4, participants were asked questions of the form *"Is it morally correct to push the man off of the bridge to save five people, even though the man would die?"* (emphasis added). This allowed us to examine whether questions that highlight the consequences of engaging in utilitarian action affect our results. In Studies 5 and 6 participants were asked to make a choice rather than to provide a moral judgment, which was of the form *"Would you push the man off of the bridge to save five people?"* This allowed us to examine whether using a foreign language affects both moral judgment and choice. In all cases, participants responded by selecting either "Yes", "No", or "I don't understand".

Moral Dilemmas. Our experiments used two different sets of moral dilemmas. Studies 1 and 2 used the original set of dilemmas from Conway and Gawronski (2013), whereas Studies 3–6 used an updated set of dilemmas (Conway & Rosas, in preparation). Both dilemma sets were comprised of 20 scenarios and were designed to recover separate *U* and *D* parameters for each

participant, but differed in terms of the specific context or story behind the choice. This allowed us to examine the robustness of the phenomena across a range of different situations.

	N	Native	Foreign	Moral Dilemmas	Elicitation Format
Study 1	214	German	English	Original Stimuli from Conway & Gawronski (2013)	Judgment (e.g., "Is it appropriate to push the man off the bridge?")
Study 2	242	English	Spanish	Original Stimuli from Conway & Gawronski (2013)	Judgment (e.g., "Is it appropriate to push the man off the bridge?")
Study 3	195	Spanish	English	Revised Stimuli from Conway & Rosas (in preparation)	Judgment with consequences highlighted (e.g., "Is it morally correct to push the man off the bridge to save five people, even though the man would die?")
Study 4	211	German	English	Revised Stimuli from Conway & Rosas (in preparation)	Judgment with consequences highlighted (e.g., "Is it morally correct to push the man off the bridge to save five people, even though the man would die?")
Study 5	209	German	English	Revised Stimuli from Conway & Rosas (in preparation)	Choice (e.g., "Would you push the man off the bridge to save five people?")
Study 6	206	English	German	Revised Stimuli from Conway & Rosas (in preparation)	Choice (e.g., "Would you push the man off the bridge to save five people?")

Table 3: Overview of experiments.

Results

For each participant, we recovered a single *U* and *D* parameter from their choices according to equations (1) and (2). Similar to Conway and Gawronski (2013), we did not observe a reliable correlation between *U* and *D* scores in any of our experiments with the exception of Experiment 3 ($r = –.20$, $p = .007$; see Table 4 for more detail). When restricting our analysis to only incongruent trials (similar to traditional measures of utilitarianism), responses in all experiments are positively correlated with *U* and negatively correlated with *D* (all *p*-values < .001). This provides empirical support that both *U* and *D* scores predict traditional measures of utilitarianism, but are independent of each other.

	U & D		**U & Traditional U**		**D & Traditional U**	
Study	***r***	***p-value***	***R***	***p-value***	***R***	***p-value***
1	-0.052	.449	0.724	<.001***	-0.697	<.001***
2	0.065	.317	0.632	<.001***	-0.698	<.001***
3	-0.193	.007**	0.765	<.001***	-0.696	<.001***
4	-0.101	.142	0.742	<.001***	-0.669	<.001***
5	-0.011	.870	0.664	<.001***	-0.670	<.001***
6	-0.033	.637	0.704	<.001***	-0.650	<.001***

Table 4: Correlations among U, D, and Traditional U parameters. *p < 0.05, **p < 0.01, ***p < 0.001

Primary Analysis. For all results reported below, we meta-analyze data across studies using a random-effects model (Lipsey & Wilson, 2001). Treating "Study" instead as a fixed effect returns even stronger results.

Overall, our results are consistent with the blunted deontology account. As illustrated in

Table 5, which reports the raw means for each experiment, we consistently observe lower *D* scores for participants in the foreign language condition compared to participants in the native language condition, with a combined effect size of $d = 0.24$ ($p < .001$). The standardized mean differences for all six experiments are displayed graphically in Figure 1. While the magnitude of the language effect varied somewhat from study to study, those using the native language responded more deontologically than foreign language users across all six experiments.

While we found consistent evidence of blunted deontology, we did find support for the heightened utilitarianism account. As illustrated in Figure 1 and displayed in Table 5, in none of the six studies do we find a reliable increase in utilitarianism as a function of foreign language use, and in three experiments we observe a reliable *negative* effect on utilitarianism. Across our experiments, participants in the foreign language condition had lower *U* scores than participants in the native language condition (combined effect size: $d = .25$, $p = .022$). These findings are in direct opposition to the claim that foreign language use increases utilitarianism.

We also examined how foreign language use affected responses when restricting our analysis to only incongruent trials (similar to traditional measures of utilitarianism). As illustrated in Table 5, we do not observe significant results for any of the experiments nor when results are combined, with a combined effect size of $d = 0.010$ ($p = .889$). This null effect is inconsistent with previous findings (e.g. Costa et al., 2014b; Corey et. al., 2017; Geipel et al., 2015a), but we note that our dilemmas differed from those used in earlier studies. Previous studies also did not directly juxtapose incongruent and congruent dilemmas, which could have led to contrast effects in the current study.

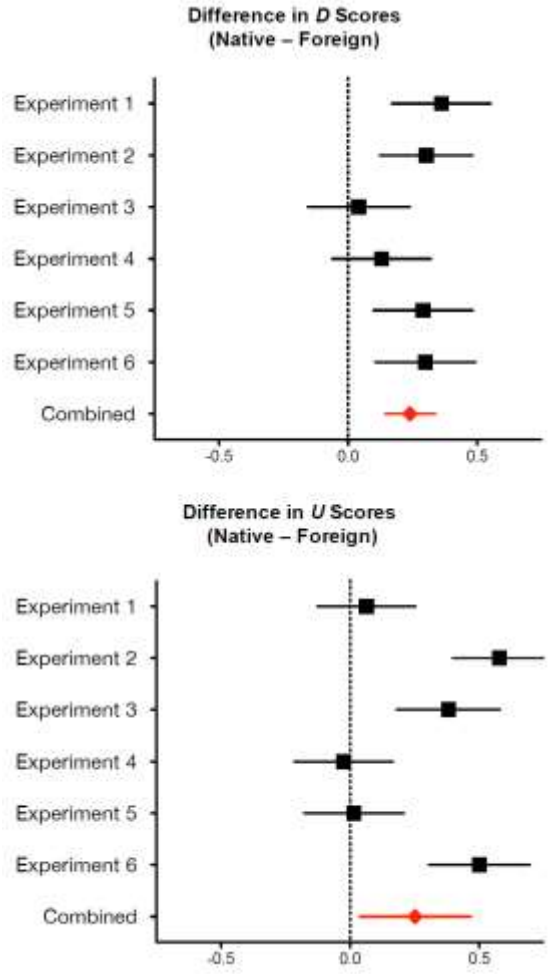

Figure 1: Forest plots of results (Studies 1–6). The graph above plots the standardized mean differences (i.e., Cohen's *d*) in *U* and *D* scores between native and foreign languages. Positive numbers indicate higher *U/D* scores in the native language condition, relative to the foreign language condition. Error bars represent 95% confidence intervals.

	Deontological Responses (*D*)			Utilitarian Responses (*U*)			Traditional Utilitarianism		
	Foreign Language	**Native Language**	***p*-value**	**Foreign Language**	**Native Language**	***p*-value**	**Foreign Language**	**Native Language**	***p*-value**
Study 1	0.65 (0.21)	0.72 (0.18)	0.009**	0.33 (0.20)	0.34 (0.19)	0.649	0.56 (0.19)	0.52 (0.19)	0.144
Study 2	0.61 (0.19)	0.67 (0.20)	0.019*	0.26 (0.18)	0.36 (0.18)	< .001***	0.55 (0.17)	0.57 (0.17)	0.31
Study 3	0.75 (0.25)	0.76 (0.29)	0.769	0.28 (0.24)	0.37 (0.25)	0.009**	0.45 (0.25)	0.50 (.25)	0.169
Study 4	0.80 (0.24)	0.83 (0.25)	0.345	0.24 (0.27)	0.23 (0.24)	0.857	0.37 (0.26)	0.36 (0.28)	0.735
Study 5	0.70 (0.30)	0.78 (0.26)	0.038*	0.32 (0.27)	0.32 (0.23)	0.915	0.53 (0.26)	0.47 (0.24)	0.107
Study 6	0.70 (0.26)	0.78 (0.29)	0.033*	0.31 (0.24)	0.43 (0.26)	< .001***	0.52 (0.23)	0.54 (0.27)	0.661
Combined results			< .001***			0.022*			0.889

Table 5: Means (standard deviations) for U, D, and Traditional U parameters. P-values reflect the results of t-tests comparing responses between native and foreign language groups for each of the six experiments as well as the combined analysis. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$

Discussion

Past research has shown that people are more willing to sacrifice one person to save five when they use their foreign language rather than their native tongue. Here we investigate *why* foreign language use affects moral choice in this way. In particular, we explore three potential accounts of the phenomenon: the blunted deontology, the heightened utilitarianism, and a hybrid account. On the one hand, the difficulty of using a foreign language might slow people down and increase deliberation, amplifying utilitarian considerations of maximizing welfare. On the other hand, use 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 find clear evidence that foreign language use blunts deontological responding.

We also consistently did not find support for the heightened utilitarianism account. In fact, in three experiments, using a foreign language reliably *decreased* utilitarian responding. It is unclear why we observe this effect, as the pattern was not consistently elicited by particular stimuli. It might have resulted from an increase in cognitive load. Indeed, using a foreign language could be cognitively demanding especially for less proficient speakers (Plass, Chun, Mayer, & Leutner, 2003), and cognitively demanding tasks debilitate utilitarian responding (Greene et al., 2008). So although using a foreign language may have led to lower *D* scores by stunting System-1 processing, foreign language use may have also induced lower *U* scores by increasing cognitive load among our less proficient foreign-language participants (we note that this explanation is independent of the mechanisms we explore in this paper). Because participants rated their foreign-language proficiency, we were able to test this explanation. We find tentative support for this theory in two of three experiments where we observe a foreign

language reduction in *U* scores — in those studies, we find that lower levels of foreign language proficiency are associated with larger reductions in utilitarian responding in the foreign language condition (see Proficiency Effects in Appendix). While these findings cannot explain why foreign-language proficiency affects *U* scores in some populations and not others, they suggest that cognitive load may play an independent role on the moral foreign language effect and should be examined in future research.

Another surprising result is that the findings do not replicate the moral foreign language effect when using a "traditional" measure of utilitarian responding. When looking only at "incongruent" moral dilemmas, which directly pit utilitarian and deontological concerns against each other, foreign language users were no more willing than native language users to endorse sacrificial harm. While this appears at odds with prior research, our stimuli were different and participants in our study considered multiple dilemmas (rather than a single dilemma) that involved direct contrasts between congruent and incongruent versions. Therefore, multiple variables could account for this discrepancy.

The dilemma whether to sacrifice one life in order to save many is consequential. On the one hand, sacrificing a life is often morally prohibited, and on the other, the lives of five people are surely worth saving. It is surprising that such a choice would affected by language. Our experiments suggest that people are more utilitarian in a foreign language not because they think more, but because they feel less.

STUDY 7:

RISK AVERSION AND PHYSIOLOGY

Studies 1-6 presented data suggesting that people using a foreign language make different moral choices because they are less emotional rather than because they are more deliberative than people using their native tongue. The goals of Study 7 are two-fold. First, we sought to examine the question of whether using a foreign language reduces emotion or increases deliberation in the domain of risk. In order to address this question, we explored whether using a foreign language would lead to a general increase in risk taking, as might be expected by a reduction in an emotional aversion to risk, or more strategic risk taking as might be expected by an increase in deliberation. Second, we investigated whether a reduction in emotion would result in different physiological responses when utilizing a foreign language compared to a native tongue. This was done by collecting ECG data while completing a gambling task in either the native or foreign language.

Many daily activities involve some degree of risk, from driving to work to investing in the stock market. However, despite extensive experience with risk we do not seem to manage it optimally, even in our domain of expertise (e.g. Haigh & List, 2005). Sometimes we play it too safe and miss out on a good opportunity (e.g. Shiv et al., 2005; Rabin & Thaler, 2001). Other times we take risks when we should have walked away. This is especially the case when walking away entails accepting a loss – we hate the idea of a sure loss and often risk losing even more in an attempt to avoid it (Kahneman & Tversky, 1979). Such emotional reactions may be in conflict with our more reflective cost-benefit calculations (Kahneman, 2003). Here we explore whether and how using a foreign language impacts the way people make risky choices.

As noted in the introduction, Keysar et al. (2012) and Costa et al. (2014a) found that people using a foreign language were more likely to accept advantageous bets than those using a native tongue. However, the risky decisions in these studies had higher expected values than the safe alternatives, making it unclear whether foreign language users were taking strategic risks or if they were just taking more risks in general.

One could imagine two different accounts for the existing findings that people take more risks in a foreign language. The first is the *Indiscriminant Risk Taking* account. This account predicts that people using a foreign language would take more risks in general, regardless of whether those risks are beneficial or not. Such a result may suggest a dominant role for reduced System-1 processing when using a foreign language. Research on patients with damage to emotion-related areas such as the orbitofrontal cortex (OFC) have demonstrated such a general increase in risk-seeking behavior (e.g. Berlin, 2004). Damasio's (1994) seminal work on the role of emotions in decision making have similarly demonstrated that damage to the ventromedial prefrontal cortex leads to less strategic risk-taking behavior which is correlated with reduced physiological arousal in anticipation of risky choices. In other words, using a foreign language may blunt the fear response to risk, leading to increased risk-taking behavior that is not sensitive to expected value.

The second possibility is what we term the *Strategic Risk Taking* account. If the psychological distance or disfluency of using a foreign language leads to more deliberative thinking, it may result in more calculated risk-taking. The maximization of expected value has been shown to be correlated with cognitive capacity, suggesting that increased deliberation would promote strategic risk taking (Stanovich & West, 2008; Benjamin & Shapiro, 2005; Frederick, 2005). In addition to being more sensitive to good vs. bad bets when using a foreign

language, they may become less sensitive to the domain of the gamble. In the domain of losses, people using a foreign language may be less impacted by the fear of sure losses that sometimes encourages people to take risks that are not beneficial. In the domain of gains, people using a foreign language may be less afraid of the possibility of losing a sure gain that sometimes prevents people from taking beneficial risks. This account would predict that people using a foreign language would take more risks when they are beneficial, but will take fewer risks when they are less beneficial than the safe alternative. To examine this question, participants were presented with a series of gambles that varied in both expected value and domain (gains vs. losses) in either their native or foreign language.

In addition to behavioral measures, Study 7 attempted to shed light on the processes underlying the FLE by collecting electrocardiogram measures to examine the effects of language on heart rate and heart rate variability (HRV). HRV refers to the variability in the timing of successive heartbeats and is taken as an index of the balance between sympathetic activity, which increases heart rate, and parasympathetic activity, which decreases heart rate. HRV at the high, respiratory frequency band (0.15-0.40 Hz) primarily represents parasympathetic, vagal activity, and has been taken as a marker of how flexibly and effectively the body responds to physical and environmental demands (Thayer et al., 2012). Greater levels of HRV are indicative of such flexibility and has been associated with greater emotional regulation (e.g. Appelhans & Luecken, 2006; Thayer & Lane, 2009). Increased emotional regulation could lead to more strategic risk taking as illustrated by Bhatt & colleagues (2015) who found that individuals with high HRV were less likely to experience negative outcomes when presented with the Balloon Analog Risk Task (Lejuez et al., 2002) due to greater inhibition of detrimental risky behavior. Individuals with higher resting HRV have also been shown to be less susceptible to framing

effects (i.e. descriptions of gains versus losses; Sütterlin et al., 2010). This could result in risk taking that is more consistent across the gain and loss domains. In addition to HRV, heart rate itself may be associated with risk taking. For example, in line with the under-arousal theory of impulsivity (Barratt & Patton, 1983), Mathias & Stanford (2003) found that healthy males scoring high on trait impulsivity had significantly slower heart rates at rest but greater relative reactivity when presented with a challenging activity compared to low-impulsive males. Depending on the pattern of physiological responses observed when using a foreign language, we may be able to better understand the processes underlying any behavioral differences observed.

In addition to overall patterns of physiological arousal, we will examine the relative reactivity to risks taken in the domain of gains vs. losses. If participants have stronger physiological responses to losses relative to comparable gains (e.g. increased heart rate), this may be taken as an indication of loss aversion (Kahneman & Tversky, 1979). In regard to heart rate variability, there are a number of potential patterns that could be expected. One possibility is that there would be a relative reduction in HRV for losses compared to gains such that losses signal a greater threat, prompting preparation for action by inhibiting the parasympathetic system. Indeed, past research has shown that stressful tasks in the lab can result in a decrease in HRV (Hansen, Johnsen & Thayer, 2003). Alternatively, there could be an increase in HRV for the loss domain, as past research has demonstrated that successful emotional regulation can result in increased parasympathetic activity (Butler, Wilheim & Gross, 2006). In either case, if using a foreign language reduces emotional reactivity, and thus loss aversion, we would expect less of a difference between loss and gain domains in a foreign language relative to a native

tongue. Such a finding would provide support for the notion that increased risk taking when using a foreign language may be related to decreased loss aversion.

Methods

Translation practices and exclusion criteria were the same as in previous studies (see Studies 1-6).

Participants and Procedure

Our sample consisted of 97 native English speakers who spoke Spanish as a foreign language. Data from an additional 24 participants were excluded from our analysis. Of these, 4 were excluded because they grew up speaking Spanish at home, 5 were not native English speakers, and 15 failed two comprehension questions which asked participants to identify the amounts that could be won or lost in different types of bets. Demographic information for the remaining participants can be found in Table 6.

Native	Foreign	Female	Age	AOA	Proficiency-Native	Proficiency-Foreign
English	Spanish	60%	20	12	6.96	4.85

Table 6: Demographic information. "AOA" is the age of foreign language acquisition.

All participants were residing in Chicago, IL at the time of the experiment. Participants were randomly assigned to either the native English condition (N=50) or the foreign Spanish condition (N=47).

The participants were endowed with $45 in small bills and coins. They were then informed that they could keep any money left at the end of the game and completed four blocks,

each with 15 gambles. Two blocks were in the loss domain. In these blocks, participants could either pay $1.00, or else gamble by flipping a coin. If they won the coin toss, they would not lose any money, but if they lost the toss, they would lose some amount of money depending on the condition. In the Loss-Good condition, losing the coin flip resulted in a loss of $1.50, which gave the gamble an expected value of -$0.75. This gamble then has a higher expected value than paying $1.00. In the Loss-Bad condition, losing the coin flip resulted in a loss of $2.50, giving the gamble an expected value of -$1.25. This has a lower expected value than paying $1.00.

The other two blocks were in the gain domain. Participants needed to decide whether to keep $1.00 or else gamble by flipping a coin. If they lost, they lost the dollar. If they won, they either got $1.50 (Bad) or $2.50 (Good). Here again, one of the conditions involved a gamble that was more beneficial than taking the safe option, while the other did not. In this way, we had four conditions for each participant – 2 for which gambling was beneficial and 2 for which it was not, as well as 2 in the domain of gains and 2 in the domain of losses. The four blocks were presented in a fixed order: Loss-Bad, Loss-Good, Gain-Bad, and Gain-Good. Note that unlike in previous experiments, bet type was varied within-subject rather than between.

Before beginning the task, participants read written instructions describing the game without specific information regarding the pay offs. Once the instructions were understood, the experimenter showed participants a card that had the specific parameters of the block written on it (see Figure 2 for example). The 15 rounds for that block then commenced. After 15 rounds, the experimenter showed the card for the next block and continued. Upon completion of the gambling task, the participant was asked to confirm how much money was remaining and told that they would be paid that amount at the end of the experiment.

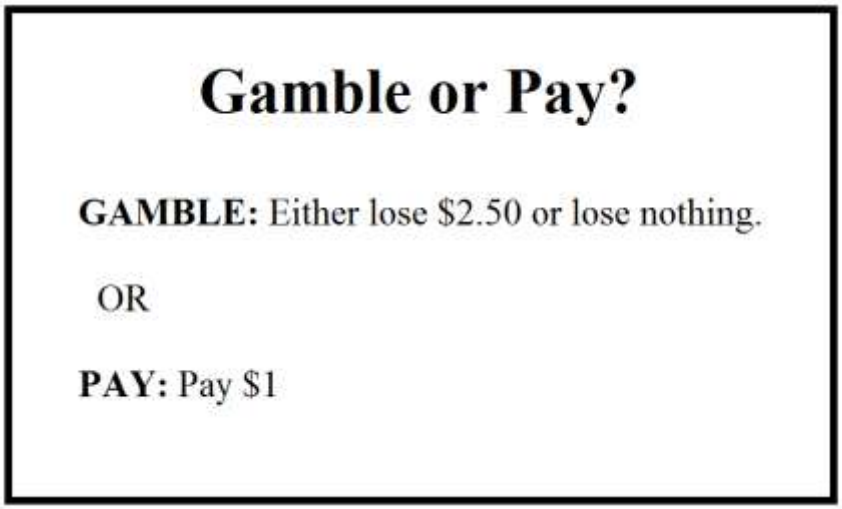

Figure 2: Example stimuli of card shown to participant that outlined the outcomes of one of the betting blocks, in this case, Loss-Bad.

If a foreign language leads to a general increase in risk taking behavior, we would expect those using a foreign language to accept more bets than those using their native tongue. This increase might be especially pronounced in the domain of gains in which people using their native language tend to be more risk averse (Kahneman & Tversky, 1979). On the other hand, if using a foreign language leads to a strategic approach to risk, we would expect different patterns of results depending on whether the bets are beneficial or not. More specifically, we would expect that in the domain of gains, in which people tend to be risk averse, that those using a foreign language should accept more Good bets, but not more Bad bets. In the domain of losses, in which people tend to be risk seeking, we would expect those using a foreign language to accept fewer Bad bets, but not fewer Good ones.

Results

Gambling Decisions

Figure 3 shows the average percentage of bets taken. To assess the effect of language (native vs. foreign), valence (good vs. bad) and domain (gain vs. loss) on risk-taking, we

performed a generalized linear mixed-effects analysis. The response variable was whether each of the 60 bets was taken or not. Language was entered as a fixed effect, and Participant, Bet (1-20), Valence, and Domain, which varied within-subject, were entered as random effects with random intercepts and slopes without random correlations. There was no main effect of Language, suggesting that using a foreign language did not lead to a general increase or decrease of risk taking (χ^2(1, N=97) = 0.117, *p=.733; b= -0.09, SE=0.28*). There was a significant main effect of Valence such that people were more likely to accept Good bets than Bad bets (χ^2(1, N=97) = 60.85, *p<.001; b= 1.97, SE=0.22*), but this did not interact with Language (χ^2(1, N=97) = 0.41, *p=.524; b= -0.21, SE=0.43*). Additionally, consistent with prospect theory, people were significantly more likely to accept bets in the domain of losses than in the domain of gains (χ^2(1, N=97) = 17.67, *p<.001; b= -0.97, SE=0.22*), but this did not interact with Language (χ^2(1, N=97) = 0.58, *p=.448; b= -0.33, SE=0.43*). We find no evidence that people are more or less sensitive to valence depending on domain (Valence x Domain interaction; χ^2(1, N=97) = 1.78, *p=.182; b= 0.22, SE=0.16*), nor was there a Language x Valence x Domain interaction (χ^2(1, N=97) = 0.24, *p=.624; b= -0.16, SE=0.31*). Contrary to past results in previous literature, we find no evidence that using a foreign language leads to more or less risk taking.

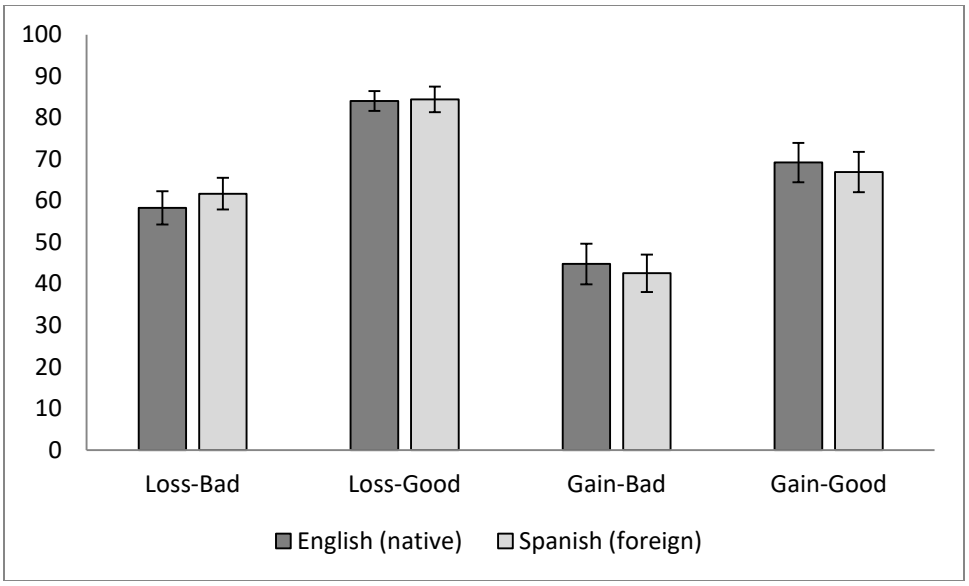

Figure 3: Gambling results. The average percentage of bets accepted out of 15 for each of the four bet conditions. Error bars represent standard errors.

Physiological Measures

A subset of participants contributed usable ECG data from which we derived heart rate and heart rate variability measures (N=22 for native and 20 for foreign)[1]. Data was collected using a standard lead II configuration. The data was then corrected for artifacts and analyzed using Mindware 3.1.2. Respiratory sinus arrhythmia (RSA) was taken as a measure of high frequency HRV. RSA refers to naturally occurring changes in heart rate during breathing such that inspiration inhibits vagal activity, resulting in a faster heart rate, while exhalation resumes vagal activity, reducing heart rate. Given that RSA emerges from vagal activity, it can be taken as a measure of parasympathetic control (Bernston et al., 1997). This spectral frequency measure of HRV is derived using fast Fourier transform and integrated over the respiratory frequency bad (0.12-0.40 Hz). The data was preprocessed in 60 second segments and then averaged within each of the gambling blocks (Bad-Loss, Bad-Gain, Good-Loss, and Good-Gain).

[1] Unfortunately, data from an additional 42 participants were not usable due to technical issues resulting in the absence of critical timing information.

We separately entered heart rate and RSA into a repeated-measure ANOVA with each of the four gambling blocks as within-subject variables and Language and Gender as between subject variables[2]. We additionally examined the effects of Language, Domain, Valence and Gender on each of the measures at baseline, which was the time between the start of ECG data acquisition and the beginning of the gambling task[3].

Heart Rate

As can be seen in Figure 4 and Table 7, no effects of Language, Domain, Valence, Gender or any interactions were significantly predictive of heart rate during the gambling task. Additionally, no effects of language ($F(1,33)=0.12$, $p=.736$, $\eta^2=.004$), gender ($F(1,33)=0.57$, $p=.455$, $\eta^2=.02$), or their interaction ($F(1,33)=0.01$, $p=.923$, $\eta^2<.001$) were found for resting heart rate prior to beginning the task. To examine the relationship between heart rate and overall gambling behavior, we separately regressed language, resting heart rate, and gambling heart rate on the average proportion of gambles accepted, the difference in proportion of gambles between good and bad blocks, as well as between gain and loss blocks. As can be seen from Table 8, none of these variables were significantly predictive of gambling. Next, we regressed the difference in heart rate between good and bad blocks on the difference in gambling between good and bad blocks, and found this effect was also non-significant ($b=.01$, SE=.02, $p=.75$). Lastly, we regressed the difference in heart rate between gain and loss blocks on the difference in gambling for these blocks and found no effect ($b=.02$, SE=.02, $p=.37$).

[2] Participant gender did not interact with language for sensitivity to expected value or overall risk taking, but was included as a variable given extensive prior research suggesting gender affects risk taking (e.g. Byrnes, Miller & Schafer, 1999). See Appendix for full results of gender effects.

[3] It should be noted that this baseline period was not standardized and thus varied in duration across subjects. Data was derived from 60 second epochs and then averaged.

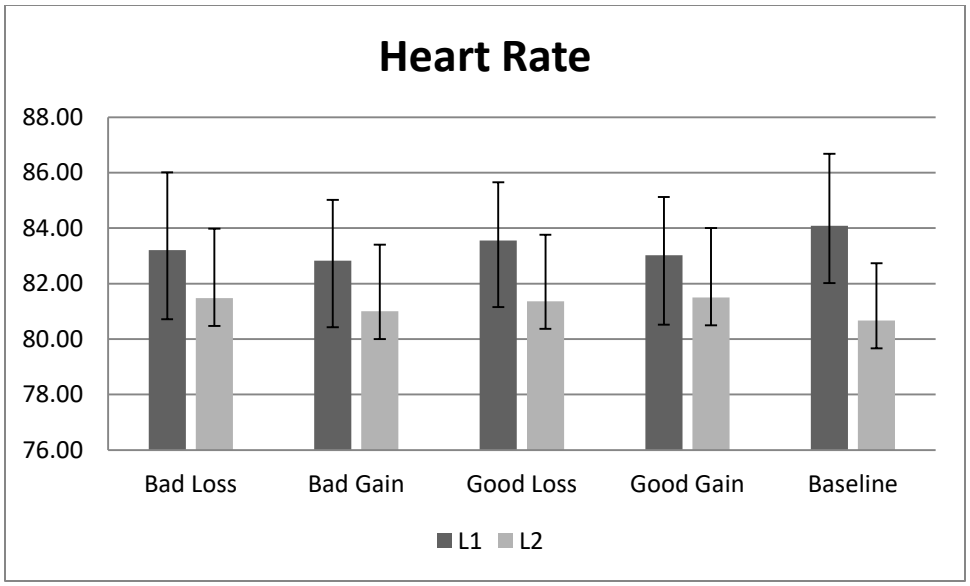

Figure 4: Average heart rate. Error bars represent standard errors.

Respiratory Sinus Arrhythmia

The same analyses conducted for heart rate were run for RSA, with comparable results. As can be seen in Figure 5 and Table 7, no effects of Language, Domain, Valence, Gender or any interactions were significantly predictive of RSA during the gambling task. Additionally, no effects of language ($F(1,33)=0.09$, $p=.765$, $\eta^2=.003$), gender ($F(1,33)=1.04$, $p=.315$, $\eta^2=.03$), or their interaction ($F(1,33)=0.38$, $p=.540$, $\eta^2=.01$) were found for resting RSA prior to beginning the task. As can be seen from Table 8, neither language, resting RSA, nor gambling RSA were predictive of overall gambling, sensitivity to expected value (Good-Bad gambling) or sensitivity to domain (Gain-Loss gambling). Additionally, as in heart rate, the difference in RSA between good and bad blocks did not predict gambling in good versus bad blocks (b =-.02, SE=.10), $p=.81$). Lastly, the difference in RSA between gain and loss blocks did not predict gambling in gain versus loss blocks (b =.08, SE=.15, $p=.60$).

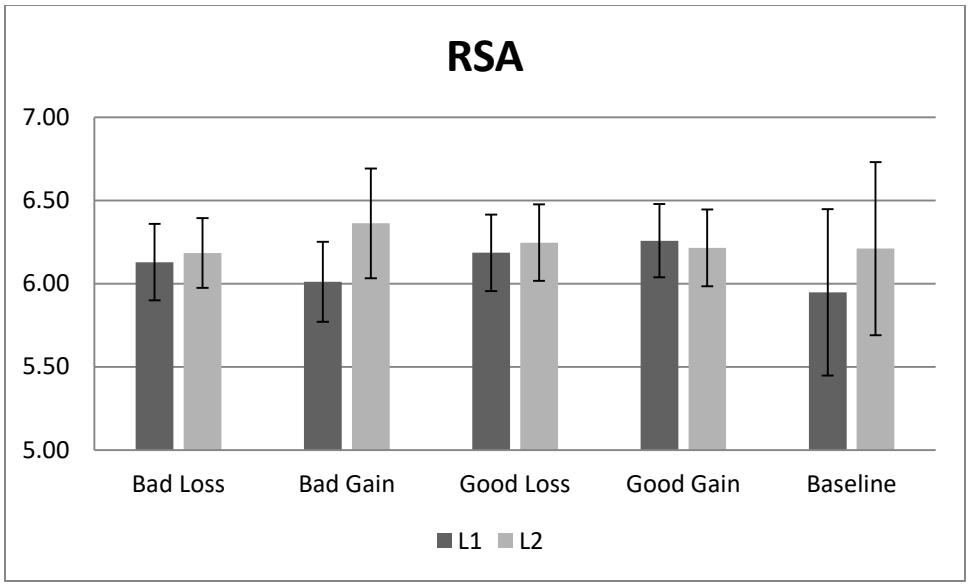

Figure 5: Average RSA. Error bars represent standard errors.

	Heart Rate		RSA	
Source	**F**	**p-value**	**F**	**p-value**
Language	0.04	0.85	0.09	0.77
Valence	1.42	0.24	2.04	0.16
Domain	0.00	0.97	0.32	0.58
Gender	1.38	0.25	0.17	0.69
Language x Valence	0.79	0.38	0.03	0.87
Language x Domain	0.13	0.72	0.22	0.64
Language x Gender	2.20	0.15	0.09	0.77
Valence x Domain	0.12	0.74	1.20	0.28
Valence x Gender	0.11	0.74	0.24	0.63
Domain x Gender	0.32	0.57	0.82	0.37
L x V x D	1.14	0.29	0.80	0.78
L x G x V	0.05	0.82	1.51	0.23
L x G x D	0.00	0.96	0.50	0.48
G x V x D	0.09	0.77	2.80	0.10
L x V x D x G	0.36	0.55	0.03	0.87

Table 7. Effects of language, valence, domain and gender on heart rate and RSA

Source	Overall Gambling	Good-Bad Gambling	Gain-Loss Gambling
Language	*b* =.04 (.10), *p*=.68	*b* =-.11 (.10), *p*=.29	*b* =.02 (.10), *p*=.86
Baseline HR	*b* =.004 (.01), *p*=.40	*b* =-.01 (.01), *p*=.21	*b* =.003 (.01), *p*=.55
Gambling HR	*b* <.001 (.01), *p*=.99	*b* =-.01 (.01), *p*=.37	*b* =-.003 (.01), *p*=.60
Baseline RSA	*b* =.01 (.05), *p*=.89	*b* =.03 (.05), *p*=.55	*b* =-.05 (.05), *p*=.26
Gambling RSA	*b* =.01 (.05), *p*=.81	*b* =.03 (.05), *p*=.56	*b* =-.04 (.04), *p*=.40

Table 8: Effects of language, heart rate and RSA on gambling behavior

Discussion

Contrary to prior research, we found no indication that using a foreign language either increased or decreased risk taking. While the reason for this discrepancy is not clear, there were some methodological differences between the current study and previous experiments that may have affected the results. For one, the present study differed from previous experiments (e.g. Keysar et al., 2012) in that it involved showing participants different types of gambles that varied both in expected value and in domain. It may be the case that the effect of language was diluted in this more complex gambling environment. As noted in the introduction, prior studies involved gambles in the domain of gains with positive expected value, whereas the present study began with the domain of losses with negative expected value. As such, it is not possible to directly compare the results from the present study with those in prior research. The fact remains however that the effect of language on risk does not appear to be especially robust across different paradigms.

One reason why the effect of language may not have emerged in the present study may have to do with a relatively low level of emotional involvement, even in the native tongue. The ECG data suggests that there was neither an increase or decrease in parasympathetic activity or heart rate during the task compared to baseline. Additionally, no differences were observed between good and bad gambles or gain and loss domains. Lastly, contrary to prior research

demonstrating that resting HRV predicts less biased behavior, such as a reduction in the framing effect (Sütterlin et al., 2010), we find no evidence that resting HRV predicts susceptibility to asymmetric gambling behavior in the domains of gains versus losses, or to gambling behavior in general. These results suggest that the task may not have been sufficiently emotional to allow for a foreign language decrease in emotion.

Lastly, the task itself may have been too repetitive and/or language-poor to evoke a foreign language effect. Unlike scenario-based tasks which have been used primarily in the moral domain (e.g. Costa et al., 2014b), the present task involved relatively little language. As such, it may be the case that foreign language effects are most robust for language-rich tasks. Indeed, such a result would be consistent with recent findings by Winskel and colleagues (2016) who discovered that people using a foreign language demonstrated reduced framing effects for language-rich tasks, but not for a relatively language-poor variant. In summary, the present results indicate that there are boundaries to the effect of language on risk, which should be systematically explored in future research.

STUDY 8:

CONSUMER DECISIONS AND PHYSIOLOGY

One man's trash is another man's treasure. This asymmetry in how individuals value goods can often lead to advantageous trades whereby a seller receives money for an unwanted item and the buyer gains something desirable in return. Such transactions become decidedly more difficult, however, when you are trying to sell treasure to a buyer who sees trash. As it happens, individuals often fall prey to this unfortunate situation due to a phenomenon known as the "endowment effect" (Thaler, 1980). The endowment effect refers to the discrepancy that emerges in the willingness to pay (WTP) to buy a good and the willingness to accept (WTA) to sell the same good. While there are numerous accounts to explain the endowment effect, a prevailing theory is that loss aversion is responsible for this WTP/WTA discrepancy (Kahneman et al., 1991). The pain of losing something looms larger than the pleasure of gaining it. As such, a seller demands more money to compensate for their considerable loss than the buyer is willing to pay to receive their modest gain. In keeping with this theory, Knutson and colleagues (2001) found that the endowment effect was predicted by increased activation in the right insula, which is associated with the anticipation of aversive events (Paulus and Stein, 2006). Surprisingly, taking acetaminophen has been shown to reduce how much sellers demand to part with their goods, presumably because it dulls the pain of loss (DeWall, Chester & White, 2015).

Given the apparent involvement of emotional processes when making buying and selling decisions, we might expect that using a foreign language may have an effect. As noted in the introduction, research has demonstrated that using a foreign language is often less emotional than

a native tongue (e.g. Pavlenko, 2005). As such, we might expect there to be an attenuation in loss aversion, which may extend to the loss of the object to sell and/or the loss of the money to buy.

If sellers using a foreign language feel less pain at parting with their goods, we would expect them to demand less money. As a result, there should be a reduction in the gap between the buyers' willingness to pay and the sellers' willingness to accept. On the other hand, if using a foreign language reduces the pain of parting with *money* (Prelec & Loewenstein, 1998), instead of goods, we might expect that buyers would be willing to pay more. This could also close the gap between buyers and sellers. Both of these processes predict a reduction in the endowment effect so long as there is an asymmetric effect of language on either the buyer or the seller (but not both). Indeed, it would make intuitive sense that the effect of language would act primarily on the party who is experiencing a loss, either the seller losing the goods or the buyer losing the money.

However, it is conceivable that the effect of language could extend to both buyers and sellers, resulting in an overall shift in price rather than a reduction in the endowment effect. For instance, if the reduction in emotion reduces the hedonic value of goods for both buyers and sellers, we might observe that individuals are both willing to accept and pay less when using a foreign language. On the other hand, if there is a reduction in the hedonic value of money for both buyers and sellers, we might expect that in addition to buyers being willing to pay more, sellers would demand more to feel satisfied with their transaction. Figure 6 illustrates these possibilities.

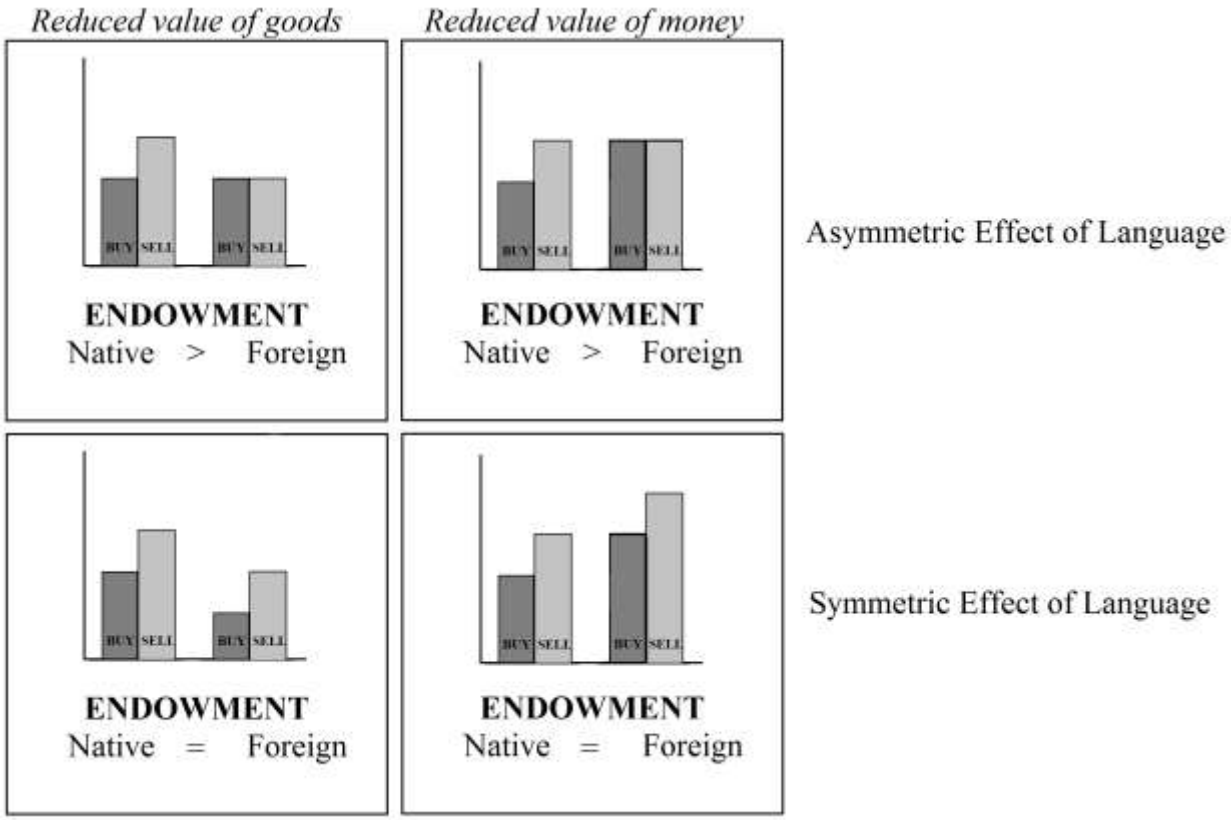

Figure 6: Potential outcomes of the effect of language

In addition to examining how language affects buying and selling decisions, we explored whether potential behavioral differences would be associated with differences in physiological activity. As noted, the size of the endowment effect has been predicted by increased insula activity (Knutson et al., 2008), which is known to be associated with the anticipation and avoidance of losses (e.g. Paulus and Stein, 2006; Samanez-Larkin et al., 2008). Tong et al. (2016) found that experience with trading reduced insula activity, which in turn mediated a reduction in the endowment effect. Prior research has also demonstrated greater amygdala activity when selling compared to buying, which appears consistent with an emotional aversion to losses (Weber et al., 2007). As a first step to explore whether using a foreign language results

in physiological states that may be indicative of a reduction in such emotional responses, we took measures of heart rate and heart rate variability (HRV) from a subset of participants.

As noted in Study 7, HRV has been linked to emotion regulation such that individuals with greater resting HRV demonstrate better regulation (Thayer & Lane, 2009). Based on this, we might expect to observe a negative relationship between resting HRV and the endowment effect such that those with better regulatory abilities, and thus higher HRV, should be least affected by loss aversion and subsequently produce a smaller endowment effect. In addition to overall HRV, we might expect to observe phasic differences in HRV across buying and selling conditions (e.g. decreased HRV for selling), which could be correlated with differences in buying and selling behavior. If using a foreign language provides emotional distance, we might expect an overall increase in resting HRV that mediates a decrease in the endowment effect. We may additionally observe less pronounced differences in HRV between buying and selling blocks when using a foreign language. Such a pattern would be consistent with the theory that using a foreign language reduces the aversion to losses and thereby leads to a reduced endowment effect.

Methods

Participants and Procedure

Our sample consisted of 128 native English speakers who spoke Spanish as a foreign language. Data from an additional five participants were excluded from our analysis. Of these, three were excluded due to insufficient comprehension of the task, one was excluded for being a non-native English speaker, and one was excluded due to technical issues resulting in the loss of data. Demographic information for the remaining participants can be found in Table 9.

Native	Foreign	Female	Age	AOA	Proficiency-Native	Proficiency-Foreign
English	Spanish	65%	21	11	6.98	4.83

Table 9: Demographic information ."AOA" is the age of foreign language acquisition.

All participants were residing in Chicago, IL at the time of the experiment. Participants were randomly assigned to complete the entire experiment either in the native tongue, English (N=64), or the foreign language, Spanish (N=64).

The procedure was based on Tong et al. (2016). In the first phase of the experiment, participants completed a lengthy filler survey in order to earn $40 which would then be used in the endowment task. After completing the filler survey and receiving their payment, participants were presented with four different items that were physically in the room (headphones, computer speakers, blender and coffee maker). They then read instructions explaining that they would complete two blocks of the experiment. In the "sell" block, they are endowed with all of the items in the room and must decide how much they would be willing to accept in order to sell each item. In the "buy" block, participants are asked to determine how much they would be willing to pay in order to buy each item. They were informed that at the end of the experiment, one of their decisions would be chosen at random to be carried out, and as such, each choice should be treated as a real decision. The order of the buying and selling blocks were counterbalanced across subjects. Within each block, participants first made a "slider decision" for each of the items. In this case, they were shown the item on the screen and were asked to use a slider in order to determine either the minimum amount they would be willing to accept to sell the item or else the maximum amount they would be willing to pay to buy the item. After making slider decisions for each of the four objects, participants were presented with a series of 124 offers (31 per item) within each block. These offers were based on the slider value such that

one offer was exactly the stated value, 15 were below and 15 were above. For each offer, participants were asked to press one of two keys to decide whether they would accept that offer ("accept" or "reject"). After responding to all 124 offers, participants took a short break and then began the second block of the experiment. For a subset of participants (N=38), we collected ECG measures throughout the endowment task after a 3 minute rest period to establish a baseline. Data collection and analysis methods were the same as those described in Study 7. After reading the instructions, the experimenter quizzed the participants on a number of key points and did not continue the study until satisfied that the participant comprehended the task.

We expected to replicate the usual endowment effect such that participants should demand more to sell an item than they are willing to pay in order to buy it. This endowment effect may be moderated by HRV such that those with higher levels should have greater emotional control and thus a reduced endowment effect. Most importantly, we hypothesized that if using a foreign language mutes emotional responding, it should likewise result in a reduction in the endowment effect and perhaps a corresponding increase in HRV.

Results

Stated Values (Slider): To examine the effect of language (native vs foreign) and task (buying vs selling) on valuation, we entered the stated values of the four items and the two tasks as within-subject variables and language as a between subject variable in a repeated-measures ANOVA. There was a significant main effect of language on the slider values ($F(1, 126)=15.13$, $p<.001$, $\eta^2=.11$). As can be seen from Figure 7, those using a foreign language required both more money to sell the objects and were willing to pay more money to buy them as well. There was additionally a significant effect of task ($F(1,126)=172.22$, $p<.001$, $\eta^2=.58$). Replicating the

usual endowment effect, people valued the objects significantly more when they were selling them than when they were buying them. Critically, we did *not* find a significant task x language interaction, suggesting that the endowment effect was not reduced when using a foreign language ($F(1,126)=.07$, $p=.8$, $\eta^2=.001$). As can be seen in the figures, some items were valued more highly than others ($F(3, 378)=7.81$, $p<.001$ $\eta^2=.06$), but this did not interact with language ($F(3,378)=1.14$, $p=.33$, $\eta^2=.009$). There was, however a significant item x task interaction such that some items evoked a stronger endowment effect than others, as can be seen in Figure 8 ($F(3, 378)=5.32$, $p=.001$, $\eta^2=.04$).[1]

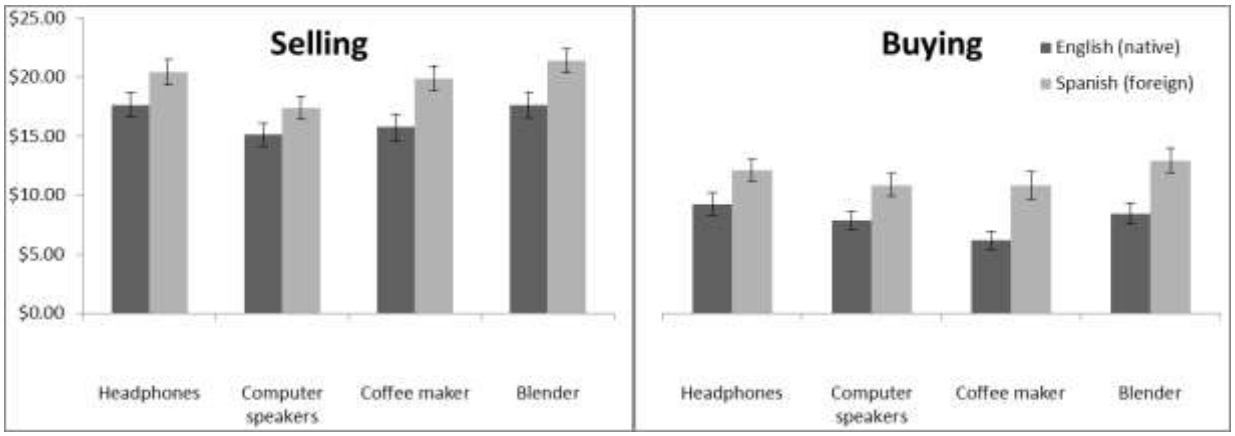

Figure 7: Slider values for buying and selling blocks

[1] The same pattern was observed when inferring values from the offer decisions. See the Appendix for a more detailed examination of behavior during the Offer section.

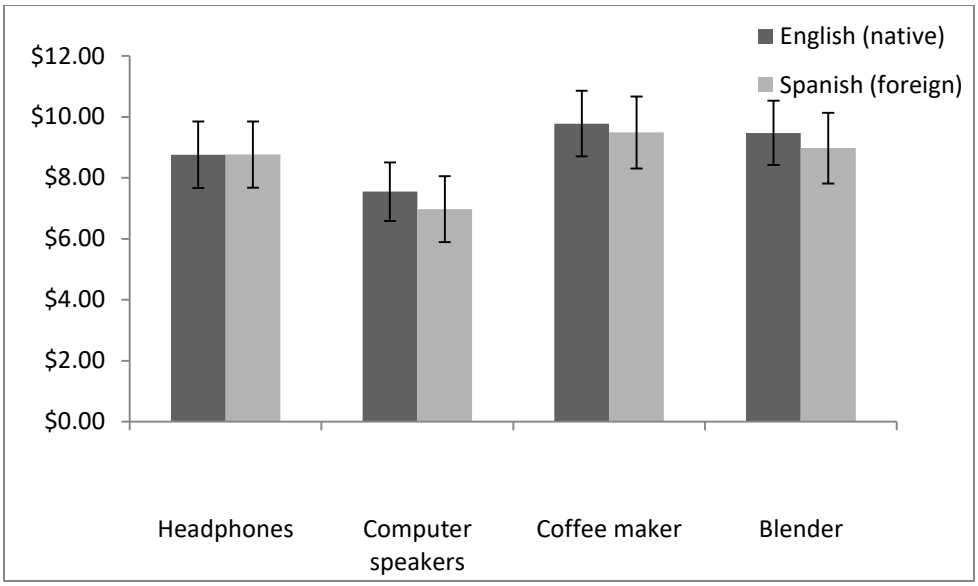

Figure 8: Endowment effect for slider values (selling minus buying)

Costs vs. Stated Values: After completing the experiment, participants were asked to estimate how much each of the items would cost. In this way, we hoped to assess evaluations of the items' values free of self-interest. While there was a significant effect of item ($F(1,122)=50.88$, $p<.001$, $\eta^2=.29$), there was no main effect of language ($F(1, 122)=.46$, $p=.499$, $\eta^2=.004$), nor an interaction ($F(1, 122)=1.70$, $p=.303$, $\eta^2=.01$). This was an intriguing result as this suggests that people using a foreign language both demand more money to sell and are willing to pay more money to buy even though they do not differ from native speakers in their objective estimation of cost. As can be seen from Figure 9, during the slider task, participants expressed that they were willing to sell for less than the estimated cost ($F(1, 122)=37.42$, $p<.001$, $\eta^2=.24$). However, this was especially the case for native speakers, resulting in a significant language x valuation interaction ($F(1, 122)=8.0$, $p=.005$, $\eta^2=.06$). Similarly, people were not willing to pay as much as the estimated cost of the item ($F(1,122)=350.54$, $p<.001$, $\eta^2=.74$), and this was especially the case of native speakers ($F(1, 122)=9.44$, $p=.003$, $\eta^2=.07$). Overall, we find that this cost-

valuation difference is greater for buying than selling ($F(1,122)=163.09$, $p<.001$, $\eta^2=.57$). Additionally, the discrepancy is significantly less for those using a foreign language ($F(1, 122)=11.57$, $p=.001$, $\eta^2=.09$) and this was not qualified by whether they were buying or selling ($F(1, 122)=.07$, $p=.789$, $\eta^2=.001$). In other words, when both buying and selling, those using a foreign language made valuations that were more aligned with the estimated cost of the item relative to those using their native tongue.

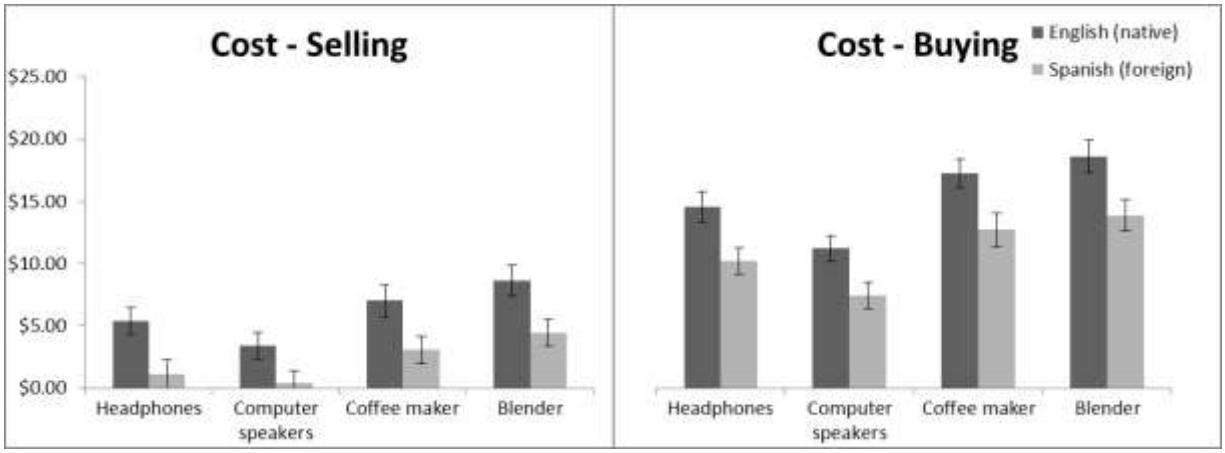

Figure 9: Difference between estimated cost and minimum stated willingness to accept (sell) and maximum willingness to pay (buy)

Physiological Measures. Data was preprocessed using Mindware 3.1.2. After removing artifacts, heart rate and RSA measures were calculated across 30 second epochs for the slider rating sections[2], and 60 second epochs for baseline and offer sections. These values were then averaged within each of the sections. Of the 38 participants with usable ECG data, four participants were excluded from the slider rating analysis as they did not have epochs long enough to calculate a meaningful score, leaving 34 participants for this section. Three

[2] Rating sections were preprocessed with shorter durations as most took less than 60 seconds.

participants were excluded from the offer analysis due to either technical errors involving timing flags or noisy data, leaving 35 participants for this section.

Heart rate: No differences were found in resting, baseline heart rate between those in the native language vs foreign language conditions (Ms = 79.43 vs 76.13, F(1, 36)=.53, p=.473, η^2=.01). In order to examine the effect of language on heart rate while participants rated the value of the items, we ran a repeated-measures ANOVA with average heart rate during buying and selling blocks as a within-subject measure, and language as a between-subject variable. As can be seen from Table 10, heart rates were not significantly different between selling and buying blocks of the slider section ("Task": Ms = 78.73 vs. 78.55, respectively). Additionally, heart rate was comparable between native and foreign speakers (Ms=77.80 vs. 79.47). Lastly, no task x language interaction was found. The same analysis was conducted on the average heart rate during the offer trials. As can be seen in Table 10, no effects of task, language or their interaction were found.

	Sliders		**Offers**	
Source	**F**	**p-value**	**F**	**p-value**
Task	0.06	0.810	0.05	0.820
Language	0.14	0.713	0.74	0.396
Task x Language	0.97	0.331	0.46	0.504

Table 10. Effects of language and task on heart rate during the slider and offer sections

Next, we examined whether there was a reliable difference between resting heart rate and the average heart rate during the task, and whether it was qualified by language. There was a significant increase in heart rate during the task relative to rest (F(1,35)=5.47, p=.025, η^2=.14). Visual inspection of the data (Figure 10) suggests that this effect was primarily driven by the

foreign language condition, though the interaction was not significant (F(1, 35)=2.50, p=.123, η^2=.07). No main effect of language was found (F(1,35)=.12, p=732, η^2=.003).

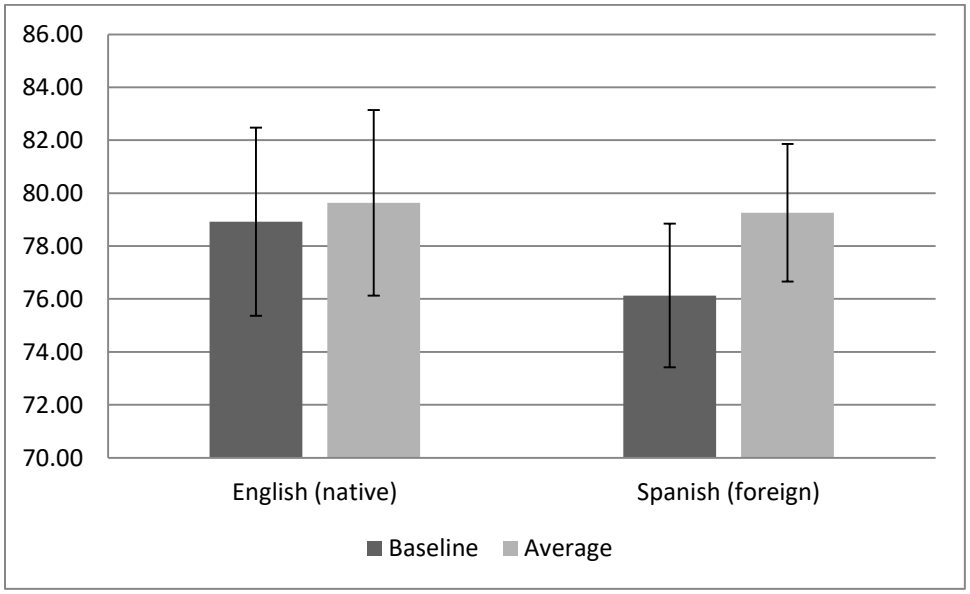

Figure 10. Average heart rate during the endowment task (slider + offer) versus baseline

RSA: No differences were found in resting RSA between those in the native language vs foreign language conditions (Ms = 6.68 vs. 6.44, F(1, 36)=.272, p=.605, η^2=.01). As can be seen from Table 11, RSA was not significantly different between selling and buying blocks of the slider section (Ms = 6.52 vs. 6.37, respectively). Additionally, RSA was comparable between native and foreign speakers (Ms=6.55 vs. 6.35). Lastly, no task x language interaction was found. The same analysis was conducted on the RSAs during the offer trials. As can be seen in Table 11, no effects of task, language or their interaction were found.

	Sliders		**Offers**	
Source	**F**	**p-value**	**F**	**p-value**
Task	0.74	0.401	1.21	0.279
Language	0.31	0.578	0.81	0.375
Task x Language	0.001	0.979	2.37	0.133

Table 11. Effects of language and task on RSA during the slider and offer sections

Next, we examined whether there was a reliable difference between resting RSA and the average RSA during the task, and whether it was qualified by language. There was a significant decrease in RSA during the task relative to rest ($F(1,35)=4.60$, $p=.039$, $\eta^2=.12$). No main effect of language was found ($F(1,35)=.22$, $p=.639$, $\eta^2=.01$) nor a language x task interaction ($F(1,35)=.55$, $p=.465$, $\eta^2=.02$).

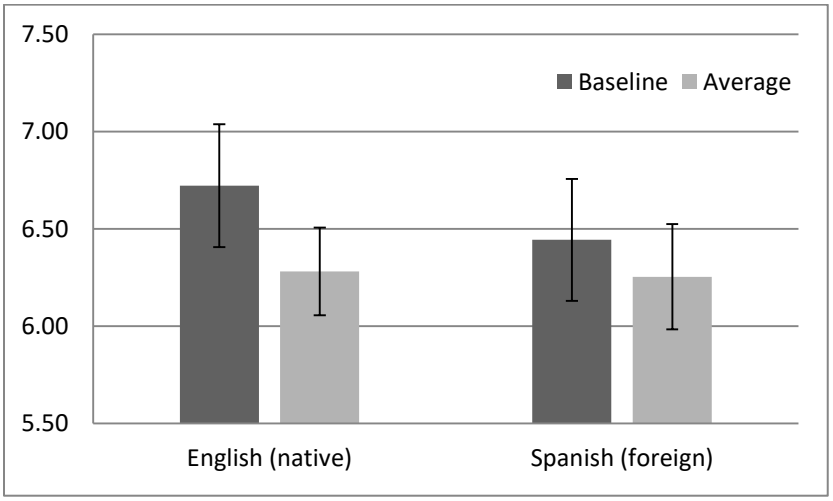

Figure 11: Average RSA during the endowment task (slider + offer) versus baseline

In summary, no effects of language or buying/selling condition were found for either heart rate or RSA. The only significant effects were between measures at rest versus during the task itself. Additionally, neither heart rate nor RSA were predictive of the size of the endowment effect. More details on this analysis can be found in the Appendix.

Discussion

We began this research by asking whether using a foreign language changes how individuals make pricing decisions when buying and selling goods. We did not find evidence that language affects the size of the endowment effect. However, we did find that using a foreign language leads to an overall increase in both buying and selling prices. As a result, prices were

more consistent with the perceived market value of the item when using a foreign language. This pattern of results could be explained by a reduction in the hedonic value of money, as described in the introduction. For example, it may be the case that the pain of paying $1 feels more like the pain of paying $0.75 when using a foreign language. As such, participants are willing to pay more to acquire an item. Likewise the satisfaction of receiving $1 feels more like the satisfaction of receiving $0.75, prompting foreign language users to demand more in order to reach the minimum hedonic threshold for feeling good about making a sale. Such an interpretation could be consistent with Weaver and Frederick's (2012) proposal that the endowment effect results from an aversion to making a "bad deal". To the extent that feelings guide such judgments (e.g. "feelings-as-information", Schwarz, 2011), a reduction in the pain of paying as well as the satisfaction of receiving money could affect the point at which individuals feel they have made a favorable deal. While such a process could explain the data, more work is needed to confirm this explanation as other possibilities remain. Two such possibilities are discussed below.

Desirability of Objects. One alternative explanation is that people using a foreign language simply desire the objects more than those using their native language. This would explain both an increase in buying and selling prices. One of the proposed explanations for the endowment effect is that buyers and sellers focus on negative and positive attributes of the objects to differing degrees, resulting in discrepant valuations (Johnson, Häubl, & Keinan, 2007). Research suggests that using a foreign language has an asymmetric effect on emotions such that negative emotions appear to be reduced more than positive (Wu & Thierry, 2012). As such, it could plausibly be the case that people using a foreign language focused on positive attributes to a relatively greater degree (in both conditions), resulting in an overall increase in liking for the objects. This does not, however appear the case. After completing the experiment, participants

were asked to rate the desirability of each of the items and no language effects were found ($F(1, 122)=1.05$, $p=.307$)[3]. As such, the present results are unlikely to be the result of an increase in liking for the objects when using a foreign language.

Anchoring on Reference Prices. Weaver & Frederick (2012) propose that endowment effects emerge, not as a result of loss aversion, but rather attention to different reference prices. When a seller considers how much she can get for her object, she will be motivated to focus on higher market prices whereas a buyer would be motivated to focus on lower reference points. In this way, the two parties diverge in what they believe is a fair price. Similarly, Simonson & Drolet (2004) propose that endowment effects emerge, in part, because sellers attend more strongly to market value than buyers, who are more influenced to subjective factors such as how well the item fits their needs. Given that the two language groups in our study demonstrated similar endowment effects, it is unlikely that using a foreign language makes people more consistent in their points of reference when making buying versus selling decisions. However, it may be the case that using a foreign language generally increases reliance on salient reference prices. The observed pattern of results suggest that those using a foreign language made both buying and selling decisions that were closer to their estimated cost of the items relative to native speakers.

It may be the case that sellers and buyers begin with estimated market prices as their anchor and then subsequently adjust (Tversky & Kahneman, 1975) to consider other factors such as their overall budget, strategy and personal preferences. Simonson & Drolet (2004) find that people are especially susceptible to even irrelevant anchors when they are uncertain about how much they value the item. Similarly, Kassam, Koslov & Mendes (2009) find that participants

[3] See the Appendix for more detailed analyses of desirability ratings as well as other factors such as gender and proficiency.

make fewer adjustments from self-generated anchors under stressful conditions. Given that using a foreign language may be more stressful and could reduce confidence in one's judgments (e.g. Geipel et al., 2015b), it may be the case that foreign language users anchored more strongly on the perceived market value of the item and did not adjust as much as native speakers for other, more subjective factors. Such a process could explain why people using a foreign language made valuations that were closer to the estimated cost of the items.

In summary, the results demonstrate that using a foreign language did not reduce the size of the endowment effect, but did lead to an overall increase in both buying and selling prices. The lack of a language effect on physiological measures precludes us from inferring that this behavioral difference is the result of a decrease in arousal or increase in emotion regulation. However, the pattern of results does raise the possibility that using a foreign language may change either the hedonic value of money or the information used to make valuations, providing exciting directions for future research into how language affects consumer behavior.

STUDY 9:

MORALITY AND MENTAL IMAGERY

From gambling to shopping, past research suggests that using a foreign language can affect the choices that people make. None of these effects, however, appear to be as dramatic or as robust as in the domain of moral decision-making. When people consider moral dilemmas such as the "Footbridge problem" (Foot, 1978; Thompson, 1985), those using a foreign language are more willing to sacrifice one person to save five (e.g. Costa et al., 2014b; Cipolletti, McFarlane, & Weissglass, 2016; Corey et. al., 2017; Geipel, Hadjichristidis, & Surian, 2015a). In Studies 1-6, we found evidence that this is because using a foreign language reduces deontological responding, which in turn may be a result of reduced activation of emotional System-1 processes. While there is substantial evidence that using a foreign language can be less emotional than a native tongue (e.g. Pavlenko, 2005; Harris et al., 2003; Dewaele, 2004, 2008; Puntoni, de Langhe, & van Osselaer, 2009), relatively little work has been done to provide a mechanistic explanation for why this is the case. Here, we investigate one factor that may help explain why using a foreign language is less emotional generally, as well as result in less deontological moral choices specifically. That is, the vividness of mental imagery when using a non-native tongue.

Imagine you are standing on a bridge overlooking a set of train tracks. You see five people tied to the track and a runaway trolley is barreling towards them. You look to your left and there is a large and heavy man next to you. The only way to stop the trolley would be to push that man off the bridge and on to the tracks, killing the man but saving the five people (Foot, 1978; Thompson, 1985). How vividly can you picture this scene? And does the vividness of your imagination affect your choice? Research suggests that it does. Amit & Greene (2012) found that

people are less likely to endorse the utilitarian action of pushing the man off the bridge if they are able to visualize him clearly in their minds. Visual interference that blocks the ability to vividly picture the scene increased the willingness to take the utilitarian, but emotionally aversive, action of sacrificing one life to save many. We propose that this link between mental imagery and moral decisions may help explain why using a foreign language increases the likelihood of taking the utilitarian action.

There are at least two reasons to think that using a foreign language could reduce the vividness of mental imagery. The first has to do with the ease with which individuals are able to access the ingredients of novel mental representations: visual memories. Imagining novel visual scenes relies on some of the same processes as imagining the past (Schacter, Addis & Buckner, 2007). Neuroimaging studies have demonstrated significant overlap in brain activity when participants are cued to recall future and past events (e.g. Okuda et. al, 2003; Szpunar, Watson, & McDermott, 2007; Addis, Wong, & Schacter, 2007). Additionally, patients with deficits in retrieving episodic memories also have trouble imagining novel scenes, further suggesting that a common machinery helps us remember events that have happened and imagine those that have not (e.g. Tulving, 1985; Hassabis, Kumaran, Vann, & Maguire, 2007). From this, there are reasons to think that using a foreign language may reduce access to memories and may thereby mute mental imagery.

Marian & Neisser (2000) found that episodic memories are *language-dependent* such that they are more easily recalled when the language in which the memory was encoded matches that of retrieval. When Russian-English bilinguals were cued to recall past events, they were significantly more likely to recall events that took place in a language environment that matched the language of the cue. Memories were also more detailed, numerous and emotional when the

language of encoding matched that of retrieval (e.g. Marian & Neisser, 2000; Schrauf & Rubin, 1998; Matsumoto & Stanny, 2006). In the case of unbalanced bilinguals, such as those used in studies examining the moral foreign language effect, it seems likely that the majority of participants' memories would have been encoded in their native tongue. As such, participants' access to their episodic memories may be reduced when imagining dilemmas in a foreign language, thereby leading to less vivid mental simulations of the scene.

Another reason why using a foreign language may reduce mental imagery is the added cognitive load of using a non-native tongue. In order to imagine scenes, individuals are required to not only recall exemplars, but must be able to flexibly recombine and reconstruct memories in order to create novel representations (Schacter et al., 2007). To the extent that such manipulations and retrieval require working memory resources, limiting such resources would likely result in some deficit. Indeed, it has been found that cognitive load reduces the vividness of mental imagery (e.g. Baddley & Andrade, 2000; Shiv & Huber, 2000). We therefore propose that reduced access to episodic memories as well as reduction in working memory resources should lead to less vivid mental imagery and subsequently different decisions while using a foreign language.

Study 9 was designed to test this hypothesis. Willingness to take a utilitarian action is decreased when potential victims of the sacrifice are described in vivid detail (Bartels, 2008) and willingness is increased when there is visual interference, thus preventing imagery (Amit & Greene, 2012). If using a foreign language similarly interferes with the generation of vivid imagery, it may explain why it increases utilitarian choice in moral dilemmas. When presented with the Footbridge dilemma, we specifically hypothesize that those using a foreign language should visualize the large man to be sacrificed less vividly relative to those using their native

tongue. This is because people afford greater weight to individuals than to groups of individuals (Redelmeier & Tversky, 1990; Kogut & Ritov, 2005) and in past research, it is the single individual to be sacrificed, rather than the five people to be saved, that is visualized most vividly and most strongly predicts participants' moral decisions (Amit & Greene, 2012).

Methods

Translation practices and exclusion criteria were the same as in previous studies (see Studies 1-6). Participants were additionally excluded if they failed to correctly describe the scenario in their native language following the experiment.

Participants and Procedure

Our sample consisted of 800 native German speakers[4]. Data from an additional 89 participants were excluded from our analysis. Of these, 52 participants were excluded for failing to correctly describe the footbridge scenario in their native tongue (German), and 37 participants were excluded because they reported German as their non-dominant language. Demographic information for the remaining participants can be found in Table 12.

Native	Foreign	Female	Age	AOA	Months Abroad	Proficiency-Native	Proficiency-Foreign
German	English	49%	38	14	3.59	6.90	5.00

Table 12: Demographic information. "AOA" is the age of foreign language acquisition and "Months Abroad" refers to the number of months spent in a country where the target foreign language is the dominant language.

Participants were randomly assigned to complete the experiment in German (n = 412) or English (n = 388). Participants first read the Footbridge dilemma. After reading the scenario,

[4] Study 9 has a large sample size because it was paired with another unrelated experiment involving 8 between-subject conditions. Language was kept constant across experiments, and did not interact with condition assignment from the other experiment for any of the four DVs in the present experiment (decision: $F(3,763) = 0.88$, $p = .45$; fat man: $F(3,763) = 0.25$, $p = .86$; five people: $F(3,763) = 0.21$, $p = .89$; overall situation: $F(3,763) = 0.34$, $p = .80$).

participants were asked "*Would you push the man?*" and responded on a 7-point scale ranging from 1 (*definitely would not push him)* to 7 (*definitely would push him*). They were also asked to rate how vividly they imagined the fat man, the five people on the track, and the overall situation. For each one, they could choose between "no image," "very vague and dim," "vague and dim," "not clear, but recognizable," "more or less clear," "very clear," and "absolutely clear image," which were coded as values 1 through 7. There was also an option for "Do not understand" and any participants selecting this option were excluded. The order of the moral decision and the imagery ratings was counterbalanced.

Results

The results replicated the finding that using a foreign language increases the willingness to endorse pushing the man compared to using a native tongue (Means = 2.41 and 2.15, respectively; $t(797) = 2.04$, $p = .04$, Cohen's $d = 0.14$). Also as predicted, participants rated their visualization of the man to be less vivid when using a foreign language than their native tongue (Means = 5.43 and 5.64, respectively; $t(797) = 2.14$, $p = .03$, Cohen's $d = 0.15$). In contrast, there was no difference in the vividness ratings of the five people who could potentially be saved between foreign and native language groups (Means = 4.72 and 4.77, respectively; $t(798) = 0.37$, $p = .71$, Cohen's $d = 0.03$), and only a marginally significant effect in the ratings of the overall situation (Means = 5.32 and 5.47, respectively; $t(797) = 1.69$, $p = .09$, Cohen's $d = 0.12$).

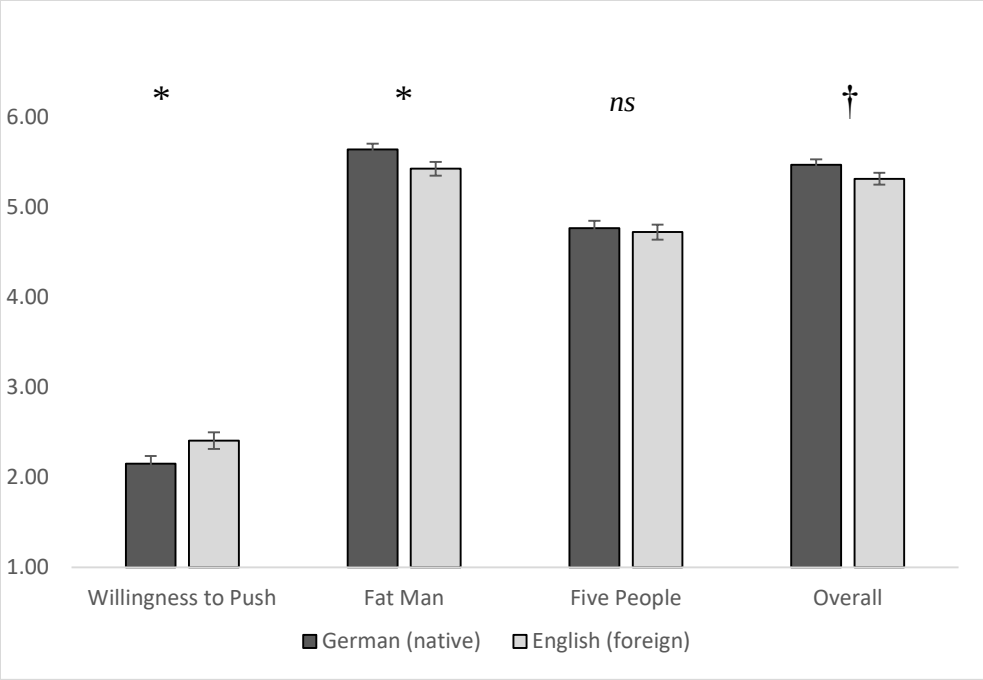

Figure 12: Mean willingness to push and vividness ratings. †p≤1, *p≤.05. Error bars display standard errors.

To evaluate the relationship between imagery and moral judgment, we regressed moral decisions onto the visualization ratings of the fat man, the five people to be rescued, and the overall situation. Replicating the findings of Amit and Greene (2012), only vividness ratings of the single potential victim were reliably associated with a willingness to push the man (b = –.129, SE = .055, p = .019). Neither vividness of the five people to be saved (b = .012, SE = .049, p = .808) nor the overall situation (b = .067, SE = .066, p = .315) significantly contributed to the willingness to push the man. We next conducted a mediation analysis to test whether the reduced vividness of the fat man explained the moral foreign language effect. Statistically controlling for vividness ratings of the fat man, the effect of language on willingness to sacrifice reduced to marginal significance (b = .24, SE = 0.13, p = .056). Using bootstrapped bias-corrected confidence interval estimates based on 5,000 resamples (MacKinnon, Lockwood & Williams, 2004; Preacher & Hayes, 2004), we find that vividness ratings of the fat man reliably mediated

the moral foreign language effect (indirect coefficient b = .018; 95% CI [.0004, .0557]). In our data, vividness ratings of the single victim statistically explained 7% of the treatment effect that foreign language use had on participants' moral judgment. Thus, visualization ratings appeared to play a small but reliable role in explaining the foreign language effect, suggesting that foreign language use affects moral judgment partly by inhibiting mental visualization of the potential victim to be sacrificed.

Discussion

The present findings are complementary to the leading explanation for why using a foreign language might increase utilitarian moral responding: that using a foreign language is less emotional than a native tongue (e.g. Pavlenko, 2005; Harris, Ayçiçegi, & Gleason, 2003). Past research has suggested that relying on more emotional or intuitive processes may encourage participants to adhere to deontological rules such as "do no harm" while more deliberative, analytical processes may favor utilitarian decisions which maximize the greater good (e.g., Cushman, 2013; Green et al., 2008). If using a foreign language is less emotional, it could explain why people seem more willing to ignore prohibitions against actively causing harm, and thus are more likely to make the utility-maximizing choice to sacrifice one person for the greater good.

Our findings are among the first to suggest a mechanistic explanation for why using a foreign language is less emotional than a native tongue. Mental imagery is known to be a powerful factor in the experience of emotion. For instance, affective responses to victims are heightened when the stimuli facilitate more vivid imagery (Dickert et al., 2011). Simply instructing participants to process unpleasant descriptions visually rather than verbally leads to increased levels of anxiety (Holmes & Matthews, 2005). Similarly, individuals who score higher on trait imagery ability are more prone to the acquisition of phobias (Dadds, Hawes, Schaefer &

Vaka, 2004). Patients with alexithymia who experience difficulty identifying emotional states also report producing less vivid mental imagery (Aleman, 205). The reduction in the vividness of mental imagery in a foreign language may thus partially explain why it is experienced as less emotional relative to a native tongue, and thereby result in different choices when presented with ethical dilemmas.

STUDY 10:

MENTAL IMAGERY PERFORMANCE

Study 9 explored whether a foreign language reduction in mental imagery could help account for the previously observed increase in utilitarian moral responding when using a non-native language. Indeed, we found that when participants were presented with the Footbridge dilemma of whether to sacrifice one person to save five (Foot, 1978; Thompson, 1985), those using a foreign tongue were both more willing to make the sacrifice and reported visualizing the person to be sacrificed less vividly relative to those using their native tongue. As noted in the introduction to Study 9, using a foreign language may reduce the vividness of mental imagery due to a number of factors relating to autobiographical memory retrieval as well as cognitive load (see Study 9 for more detail). While the results from the previous chapter seem to align with this hypothesis, one limitation is that the measure of mental imagery relied on self-report. As such, it is conceivable that the reported difference in vividness found between languages does not reflect actual differences in visualization, but rather a response bias or else subjective experiences without an objective basis.

Some research suggests that scales may be interpreted differently when using a foreign language (e.g. “the anchor-contraction effect”; De Langhe et al., 2011). Additionally, there is research suggesting that self-report measures of vividness such as the Vividness of Movement Imagery Questionnaire (VMIQ; Isaac et al., 1986) are not reliably correlated with objective imagery measures (Lequerica et al., 2002). As such, the goal of Study 10 is to assess whether using a foreign language leads to a reduction in visualization using an objective measure.

Methods

Translation practices and exclusion criteria were the same as in previous studies (see Studies 1-6).

Participants and Procedure.

Data from 307 participants residing in Beijing, China were included in the analysis. Data from an additional 17 participants were excluded from our analysis. Of these, four were excluded because they reported the target foreign language, English, to be their dominant language, and 13 were excluded because they reported growing up speaking English at home. Demographic information for the remaining participants can be found in Table 13. Participants were randomly assigned to either the native (N=152) or foreign (N=155) language conditions.

Native	Foreign	Female	Age	AOA	Months Abroad	Proficiency-Native	Proficiency-Foreign
Mandarin	English	63%	25	10	6.32	6.62	4.65

Table 13: Demographic information ."AOA" is the age of foreign language acquisition and "Months Abroad" refers to the number of months spent in a country where the target foreign language is the dominant language.

We used a mental imagery task based on a paradigm introduced by Mehta, Newcombe, & De Haan (1992). Participants were presented with three stimuli and were asked to identify the one that was least like the other two based on a given attribute. For example, a participant could be presented with the words "carrot", "mushroom" and "pen" and be asked to identify which one was least like the other two in terms of its shape, ignoring all other attributes such as size, texture and color (see Figure 13). In this case, we would expect participants to choose "mushroom". In order to answer this question accurately, participants must visualize the items. Therefore,

accuracy on this task served as a measure of how clearly individuals were able to mentally picture the items.

If the use of a foreign language reduces the ability to visualize, then native language users should outperform those using a foreign language. However, there could be an alternative explanation to such a finding. Instead of reduced visualization, foreign language users might perform worse because they are not as competent in the language. A person who is unclear about the meaning of "mushroom" would not perform well in this task regardless of mental imagery abilities. In order to control for such effects relating to language competency, we included a second task that required the same level of knowledge of the language but did not rely on visualization. Participants were presented with the same stimuli as in the shape judgment task, but selected the odd item out on the basis of its category membership or function, not on the basis of shape. Therefore, for "carrot", "mushroom", and "pen" participants should choose "pen" as the odd one out. This category task should provide a baseline for any deficit induced by the use of a foreign language that is due to language competency such as vocabulary size. We therefore would consider any additional foreign language deficit in the Shape task above and beyond that of the Category task as more directly related to reduced visualization. Participants completed two additional blocks in which the stimuli were simple line drawings that corresponded to the words from the Word blocks, one with a shape judgment and one with a category judgment. This was to ensure that any foreign language effects are indeed due to the language of the stimuli rather than to anxiety that could have been provoked by being in a foreign language environment, leading to overall worse performance.

Each participant completed all four blocks in a counter balanced order. Each block was made up of 24 trials. For each participant, the same words were used for the Shape and Category

tasks, but different stimuli were used for the Word and Picture tasks. In this way, no participant saw a word and a picture that depicted the same item. However, they did see the same word or picture for making a shape and category judgment, allowing us to observe the isolated effect of visualization for any given stimulus. The assignment of specific stimuli to Word or Picture was counterbalanced across subjects. In order to ensure comprehension, participants were shown each of the Word stimuli again at the end of the experiment and were asked to choose which line drawing matched the word from a set of three pictures. We then excluded any trial from the primary task that utilized an incorrectly matched word.

If using a foreign language reduces the vividness of mental imagery, then native language users should outperform foreign language users on the Word-Shape task. Critically, this language difference should be greater in the Shape task than in the Category task for which no mental imagery is required. This should result in a significant Language x Task interaction for word stimuli. We do not predict such an interaction for picture stimuli. While there might be a general foreign language deficit for pictures due to cognitive load or anxiety, this should be the same for the Shape and Category tasks.

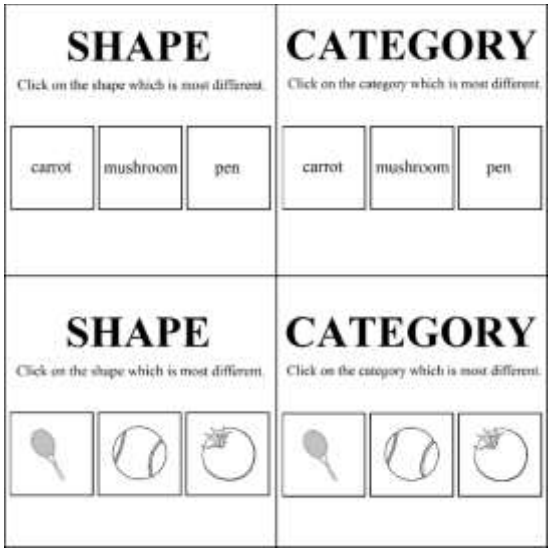

Figure 13: Example stimuli. Example trials from each of the four blocks in the English condition. Clockwise from the top left: Word-Shape, Word-Category, Picture-Category, Picture-Shape.

Results.

Word stimuli analysis. After excluding any Word trials for which participants did not correctly answer the comprehension check, we calculated an accuracy score for each of the two Word blocks. We then ran a repeated-measures ANOVA with Language (native vs. foreign) as a between-subject variable and Task (shape vs. category) as a within-subject variable. We found a significant main effect of Language such that those using a foreign language were overall less accurate than those using their native language ($F(1,305)=80.11$, $p<.001$, $\eta^2=0.208$). Additionally, there was a significant main effect of Task such that participants were more accurate on the Category task than the Shape task ($F(1, 305)=111.62$, $p<.001$, $\eta^2=0.257$). Critically, there was the predicted Language x Task interaction such that the effect of language was greater for the Shape task than the Category task ($F(1, 305)=17.39$, $p<.001$, $\eta^2=0.04$)[5]. This suggests that while there may be some costs to using a foreign language that have to do with the words or vocabulary themselves, there is an additional cost that is related to visualization.

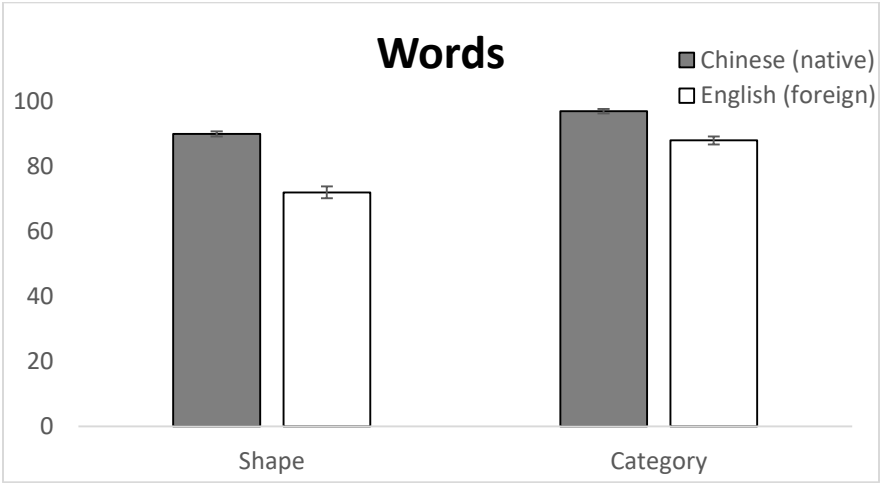

Figure 14: Average accuracy (%) for words. Error bars display standard errors.

[5] The same exact pattern of significant results were obtained when restricting the analysis to participants with perfect comprehension scores.

Picture stimuli analysis. For the two Picture blocks, we once again calculated the accuracy score for each task and then ran a repeated-measures ANOVA with Language (native vs. foreign) as a between-subject measure and Task (shape vs category) as a within-subject variable. Once again we found a small but significant main effect of Language such that those using their native language outperformed those using their foreign tongue ($F(1, 305)=7.04$, $p=.008$, $\eta^2=0.023$). This suggests that there were indeed some costs to simply being in a foreign language context that hinders performance on even non-linguistic tasks. Unlike for the Word blocks, however, we found no effect of Task ($F(1,305)=2.59$, $p=.108$, $\eta^2=0.005$), and most importantly, no Language x Task interaction ($F(1, 305)=.018$, $p=.893$, $\eta^2<.001$). The fact that there is no interaction with the picture stimuli suggests that the interaction with the word stimuli is specific to processing the language. It rules out the possibility that the interaction for words is due to a general effect of cognitive load or anxiety that a foreign language might induce.

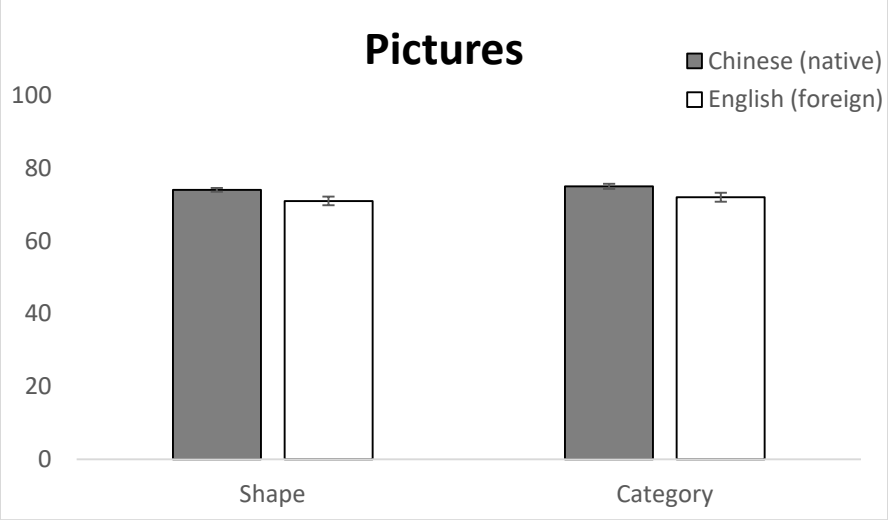

Figure 15: Average accuracy (%) for pictures. Error bars display standard errors.

In summary, the data reveals that using a foreign language is more disruptive for Shape judgments than Category judgments, but only when the stimuli are presented as words. Together, the data demonstrates that performance on tasks which require mental imagery are especially

hindered by the use of a foreign language, suggesting that the vividness of mental imagery may be diminished when using a non-native tongue.

Discussion

The results from Study 10 reveal that using a foreign language hinders performance on tasks requiring mental imagery. Together with the results from Study 9, we provide both objective measures and subjective reports that support the theory that using a foreign language leads to a reduction in the vividness of mental simulations. This reduction in visualization may help explain a number of foreign language effects including the previously discussed increase in what appears to be utilitarian moral judgment. It may also suggest a number of novel ways in which using a foreign language could affect both judgment and decision-making. For example, prior research has demonstrated that events which are easily or vividly imagined are often perceived as subjectively more likely to occur (e.g. Gregory, Cialdini, & Carpenter, 1982; Sherman et al., 1985). If using a foreign language reduces the ease or vividness of mental imagery generation, it may affect perceptions of probability and likelihood. Relatedly, it has been demonstrated that risk perceptions and willingness to take risks are affected by the intensity of mental imagery regarding potential negative consequences (Traczyk, Sobkow & Zaleskiewicz, 2015). If using a foreign tongue reduces the vividness of such visualizations, it could help explain previously observed effects such as how using a foreign language decreases risk perceptions (Hadjichristidis, Geipel, & Savadori, 2015) and can increase willingness to take risks (Keysar, Hayakawa & An, 2012; Costa et al., 2014a).

Over the last few years, there has been growing evidence that the use of a foreign language affects biases, perceptions of risk and loss, and moral choice. But there has been little

progress in understanding why this occurs. Here, we demonstrate that one reason is that the world imagined through a foreign language is less vivid than through a native tongue.

STUDIES 11 & 12:

PROBABILITY AND MENTAL IMAGERY

In previous chapters, we explored the relationship between language, mental imagery and moral judgment. Here, we examine whether a reduction in the vividness of mental imagery could help explain a different type of foreign language phenomenon. As noted in previous chapters, there is research demonstrating that using a foreign language can both increase risk taking (Keysar et al., 2012; Costa et al., 2014a) and decrease perceptions of risk (Hadjichristidis et al., 2015). There are reasons to think that this reduction in perceived risk may be partly due to a reduction in mental imagery. Prior research has demonstrated that events which are easily or vividly imagined are often perceived as subjectively more likely to occur (e.g. Carroll, 1978; Gregory, Cialdini, & Carpenter, 1982; Sherman et al., 1985; De Brigard, Szpunar & Schacter, 2013; but see Levi & Pryor, 1987). This is in line with the "availability heuristic" whereby an event is perceived as more frequent or likely when relevant instances of it easily come to mind (Tversky & Kahneman, 1973). In one experiment, Sherman and colleagues (1985) asked participants to imagine a disease that was described in terms of symptoms that were either easy to imagine (e.g. sore throat) or hard to imagine (e.g. nerve damage). Participants were then asked to rate how likely they thought it was that they would contract the disease. The authors found that participants perceived the easy-to-imagine disease to be more likely than the hard-to-imagine disease and that imagery scores predicted likelihood ratings.

If using a foreign language reduces the vividness of mental imagery and/or the accessibility of relevant instances of a given event, it may affect their judgments of perceived likelihood. This may happen in at least one of two ways. One possibility is that the reduced

vividness of imagery may lead foreign language users to perceive all events to be relatively less likely than native language users (i.e. a main effect of Language). This is a likely outcome if there is a linear relationship between ease of imagery and perceived likelihood. Another possibility is that the reduced vividness primarily affects events which are easy to imagine for native users, and not those which are hard to imagine (i.e. a Language x Difficulty interaction). This is a likely outcome if the relationship between imagery and perceived likelihood follows more of a step function such that easy-to-imagine events are perceived as generally likely while those that are hard are perceived as generally unlikely, without much sensitivity to degree within each of these categories. Either outcome could provide a potential mechanistic explanation for why using a foreign language reduces perceived risk (e.g. Hadjichristidis et al., 2015). Study 11 first examines whether ease of visualization affects the perceived likelihood of contracting a disease versus an unrelated event. Study 12 follows up by exploring the effect of visualization on the perceived probability of events involving the self versus other people.

Study 11

Methods

Translation practices and exclusion criteria were the same as in previous studies (see Studies 1-6).

Participants and Procedure.

Data from 431 native Korean speakers were included in the analysis. Data from an additional two participants were excluded from our analysis because they indicated that English, the target foreign language, was their dominant language. Demographic information for the

remaining participants can be found in Table 14. Participants were randomly assigned to either the native or foreign language conditions.

Native	Foreign	Female	Age	AOA	Proficiency-Native	Proficiency-Foreign
Korean	English	57%	23	12	6.80	4.00

Table 14: Demographic information ."AOA" is the age of foreign language acquisition

Based on Sherman et al.'s (1985) experiment, participants were asked to imagine an epidemic with five possible symptoms, and then were asked to judge the likelihood of contracting that disease. During the imagery stage, participants were presented with each of the symptoms and were asked to imagine how they would feel for 10 seconds each. Critically, participants were randomly assigned to imagine symptoms that were either easy or hard to imagine. In a pilot study, 81 MTurk participants rated the difficulty of imagining either the Hard or Easy symptoms on a scale ranging from 1 to 7 with higher scores indicating greater difficulty. As expected, the Hard symptoms were rated as more difficult to imagine relative to the Easy symptoms (Ms = 3.63 vs. 1.95, respectively; $F(1, 79)=29.96$, $p<.001$). Participants in the present study were thus randomly assigned to one of four conditions: Native-Easy (N=115), Native-Hard (N=114), Foreign-Easy (N=104), or Foreign-Hard (N=98). The Hard symptoms were as follows: "mild sense of disorientation", "nervous system damage", "liver infection", "mental confusion", and "blindness". The Easy symptoms were as follows: "low energy level", "muscle aches", "frequent severe headaches", "tiredness", and "diarrhea". After imagining the symptoms, participants were asked to rate how likely they were to contract the disease on a scale that ranged from "very unlikely" to "very likely" with higher values indicating increased likelihood. They were additionally asked to use the same scale to rate the likelihood of becoming involved in a new

romantic relationship. This latter question was included as a control, as we would expect both language and ease of imagery to affect the perceived likelihood of contracting the disease, but not an unrelated event. Both questions were displayed on the same screen with the order counterbalanced across subjects. Lastly, participants were asked to rate how difficult it was to imagine the disease symptoms on a scale ranging from "could imagine them so vividly that it felt real" to "could not imagine at all" with higher scores indicating greater difficulty.

Results

For all three measures (Disease Likelihood, Control Likelihood, Imagery Difficulty), we Z-score transformed the scale ratings prior to analysis. This was done because of an error discovered after data collection. While the scales were unnumbered from the perspective of the participant, the native language condition's likelihood scale contained 10 possible values whereas the scale for the foreign language condition contained only 9. It should be noted that this discrepancy occurred only across language conditions, and not across ease of imagery conditions. We thus attempted to address this issue by taking the normalized Z-score within each scale for each measure.

Likelihood

We first investigated the effect of Difficulty and Language on likelihood ratings. We entered the ratings for Disease and Control scenarios as repeated-measures dependent variables and Difficulty and Language as between-subject independent variables in a 2x2 mixed effects ANOVA. As predicted, there was a significant Language x Difficulty x Scenario interaction ($F(1, 429)=4.27$, $p=.039$, $\eta^2=.01$). This reflects the pattern that those using their native language were more affected by the difficulty of imagery than foreign language users when judging the

likelihood of contracting the disease (F(1, 429)=6.10, $p=.014$, $\eta^2=.01$). No such interaction was found for the control scenario (F(1, 429)=.30, $p=.584$, $\eta^2=.001$). This result is consistent with the hypothesis that people using a foreign language rely less on their mental imagery to make likelihood judgments. What was unexpected, however, was that there was a significant main effect of Difficulty for the control scenario such that individuals rated the likelihood of entering into a new relationship to be lower after being exposed to disease symptoms that were *easy* to imagine (F(1, 429)=12.58, $p<.001$, $\eta^2=.02$). This pattern is the opposite of that observed for the disease scenario and resulted in a significant Scenario x Difficulty interaction (F(1, 429)=63.42, $p<.001$, $\eta^2=.13$) as can be seen in Figure 16. One explanation for this surprising result is that people's estimates of the unrelated event were affected by their estimates of the likelihood of contracting the disease. In other words, those who felt that the disease was very unlikely gave higher ratings for the control question because it seemed *relatively* more likely.

Because we needed to calculate Z-scores within each language to deal with the scaling issue, we were unable to directly test the main effect of language (as both were transformed to have a mean of 0 and standard deviation of 1). As such, in order compare across languages, we ran the analysis with arcsine transformed proportions rather than Z-scores and did not find a significant overall difference between languages in how likely they perceived the events to be (F(1,429)=2.06, $p=.152$, $\eta^2=.01$). Therefore, the data seems most consistent with the theory that using a foreign language reduces reliance on mental imagery when making likelihood judgments, rather than an overall reduction in perceived likelihood.

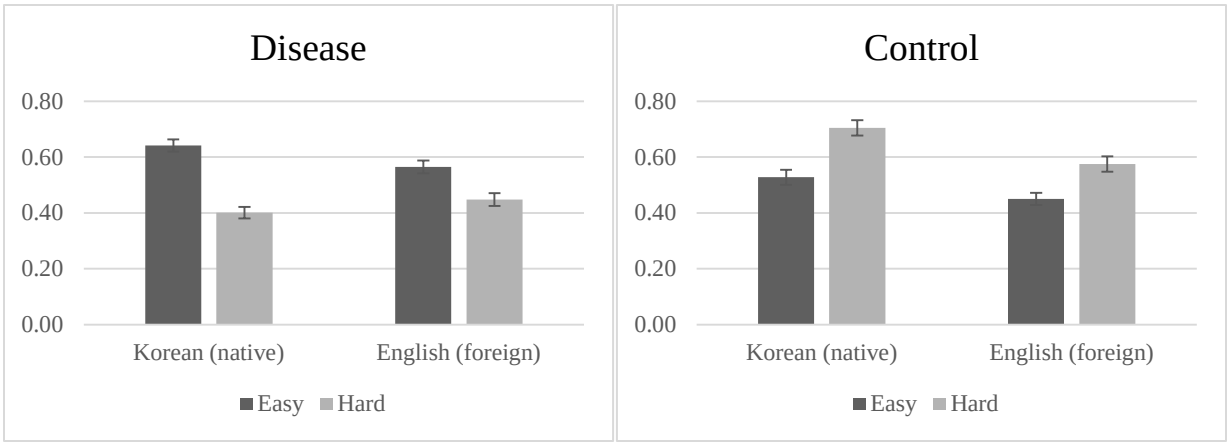

Figure 16: Likelihood ratings. The Y-axis represents perceived likelihood of the event with higher values indicating greater likelihood. Error bars represent standard errors.

Imagery

To assess whether those using a foreign language perceived the symptoms to be less imageable, we ran a univariate ANOVA with the transformed difficulty score as the dependent variable and Language and Difficulty as predictors. There was a significant main effect of Difficulty such that those in the Hard condition rated the symptoms to be more difficult to imagine than those in the Easy condition ($F(1, 429)=70.39$, $p<.001$, $\eta^2=.14$). There was, however, no main effect of Language ($F(1, 429)=2.42$, $p=.121$, $\eta^2=.01$), nor a Language x Difficulty interaction ($F(1, 429)=0.89$, $p=.346$, $\eta^2=.002$).

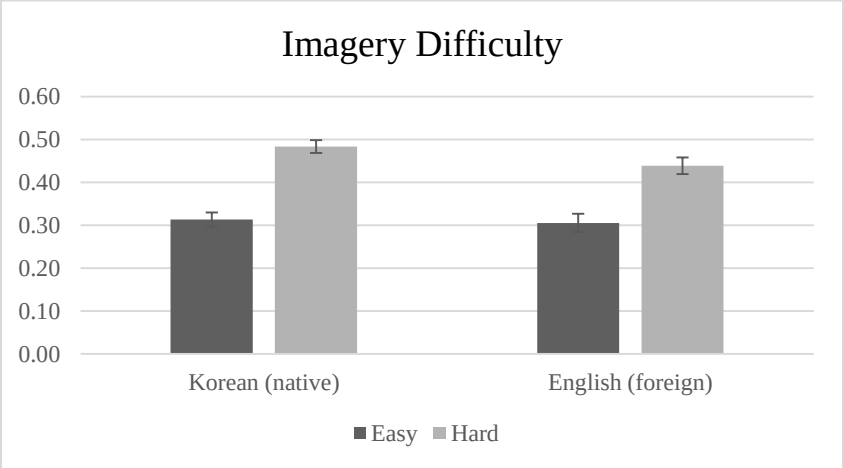

Figure 17. Difficulty imagining disease symptoms. The Y-axis represents difficulty ratings with higher values indicating greater difficulty imagining the symptoms. Error bars represent standard errors.

In summary, we found evidence that people using a foreign language were less affected by the ease of imagery when making likelihood judgments. However, we did not find evidence that using a foreign language leads to a reduction in mental imagery.

Study 12

The goal of Study 12 was to provide a conceptual replication of the previous study as well as to address a number of limitations. First, we sought to examine the robustness of the effect with a different language population of native German speakers. Second, we decided to change the control question from an estimation of an unrelated event occurring (i.e. getting involved in a new relationship) to an estimation of how likely other people are to contract the disease. Past research has demonstrated that the vividness of mental imagery has a more pronounced effect on likelihood estimations involving the self (Caruso, 2008). This change was made to address a potential problem with the previous design such that the symptoms which were more easily imagined (e.g. diarrhea) were perhaps also perceived to be more contagious than the difficult symptoms (e.g. nervous system damage), irrespective of imagery. If so, that would provide an alternative explanation for why those in the Easy condition perceived their likelihood of contracting the disease to be higher. By making the control question about other people, we can thus address this possibility, as lay-beliefs regarding the contagiousness of diseases should be the same regardless of whether participants are making judgments for themselves vs. others.

Methods

Participants and Procedure.

Data from 435 native German speakers were included in the analysis (49% female; M_{age}=30.4 years old). Data from an additional 18 participants were excluded from our analysis because they failed to correctly translate at least 4 out of 5 of the symptoms from the foreign language to the native language, as assessed by two German-English bilinguals[1]. Participants were randomly assigned to either the native or foreign language conditions.

As in Study 11, participants were randomly assigned to either the Easy or Hard imagery conditions, resulting in four different groups: Native-Easy (N=112), Native-Hard (N=116), Foreign-Easy (N=107), or Foreign-Hard (N=100). The same basic procedure and stimuli were used as in the previous study with a few exceptions. First, we changed the wording of the imagery effort question from "How difficult was it to imagine the disease symptoms?" to "How much effort was required for you to imagine the symptoms of the disease?". This was done to be more consistent with previous literature (Sherman et al., 1985). More importantly, we changed the control question from an unrelated event to assessments of likelihood for others, as noted earlier. In order to avoid demand effects whereby participants attempt to be consistent when judging likelihood for themselves vs. others, we asked the questions in slightly different ways. The Self question was the same as in Study 11 whereby participants were asked to rate how likely they were to contract the disease on a scale ranging from 'very unlikely' to 'very likely' with higher numbers indicating increased likelihood. For the Other question, participants were shown a photo of a crowded cocktail party and were told that "a person with the disease was invited to a party of 100 attendants. How many people do you think will be infected by the

[1] All analyses were run with this criteria as well as with the more conservative criteria excluding all participants who failed to accurately translate all five items (N=402 included using this criteria).

disease from the party?". In order to compare the self vs. other ratings, each value was divided by the maximum possible value resulting in a proportion.

Results

Likelihood

To investigate the effect of imagery difficulty and language on perceived likelihood, we entered the arcsine transformed likelihood proportions for the Self and Other targets as within-subject dependent variables and Language and Difficulty as between-subject predictors in a repeated-measures mixed ANOVA. As in Study 11, we observed a marginal Language x Difficulty x Target interaction ($F(1, 431)=3.23$, $p=.073$, $\eta^2=.01$). However, unlike in the previous experiment, this interaction reflects a comparable effect of image difficulty across languages when assessing likelihood for the self (contrary to Study 11), and a relative *increase* in the effect of image difficulty for the foreign language group when judging likelihood for others.

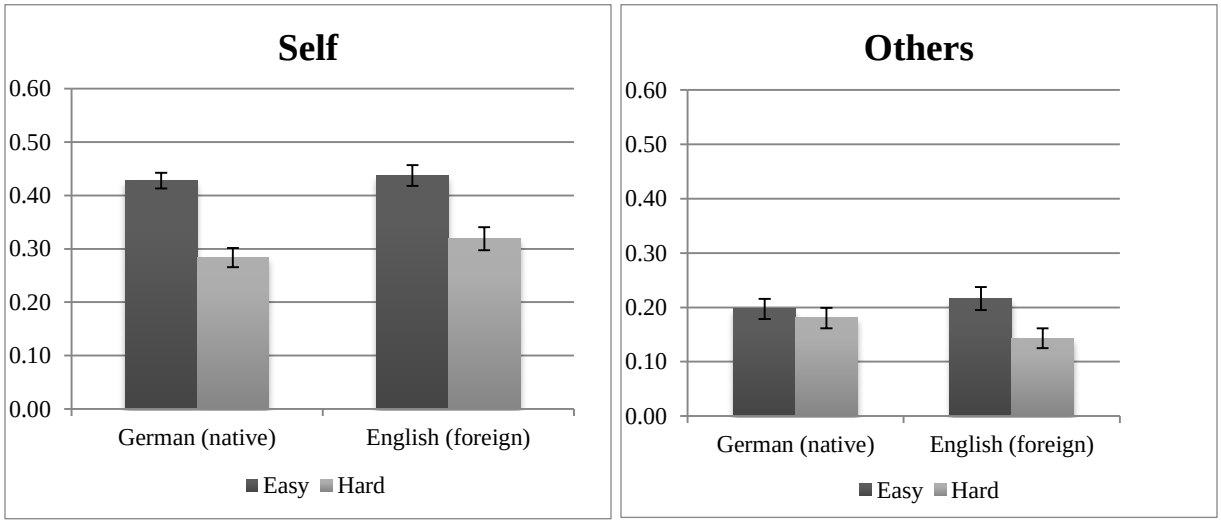

Figure 18. Likelihood ratings for the self vs. others. The Y-axis represents perceived likelihood of contracting the disease with higher numbers indicating greater likelihood. Error bars represent standard errors.

Consistent with past research, the effect of image difficulty was more pronounced when making judgments for the self relative to others, resulting in a significant Target x Difficulty interaction ($F(1, 431)=6.17$, $p=.013$, $\eta^2=.01$). Though as reflected in the previously discussed three-way interaction, this difference between self and others for the effect of imagery was greater for those using their native language relative to those using their foreign tongue. As such, while we did not replicate the previous observation that those using a foreign language are less affected by ease of imagery when making self-judgments, we do observe some language effects. Mainly, we observe here that while those using their native language are only affected by imagery when making judgments for themselves, those using their foreign language similarly utilize ease of imagery when making judgments for themselves as well as others.

Imagery

To examine the effect of language on reported difficulty of imagining the symptoms, we entered the arcsine transformed imagery score as the dependent variable and Language and Difficulty as between-subject predictors in a univariate ANOVA. As can be seen in Figure 19 and consistent with Study 11, we observe a main effect of Difficulty such that the participants in the "Hard" condition perceived the symptoms to be more effortful to imagine than those in the "Easy" condition ($F(1, 431)=71.59$, $p<.001$, $\eta^2=.14$). More importantly, however, we now observe a main effect of Language such that those using their foreign language found imagining the symptoms to be more effortful relative to those using their native tongue ($F(1, 431)=11.24$, $p=.001$, $\eta^2=.03$). This finding is consistent with the results from Studies 9 & 10. There was no Language x Difficulty interaction, suggesting that this increase in effortfulness when imagining

in a foreign language occured for both Easy and Hard symptoms ($F(1, 431)=1.03$, $p=.311$, $\eta^2=.002$)[2].

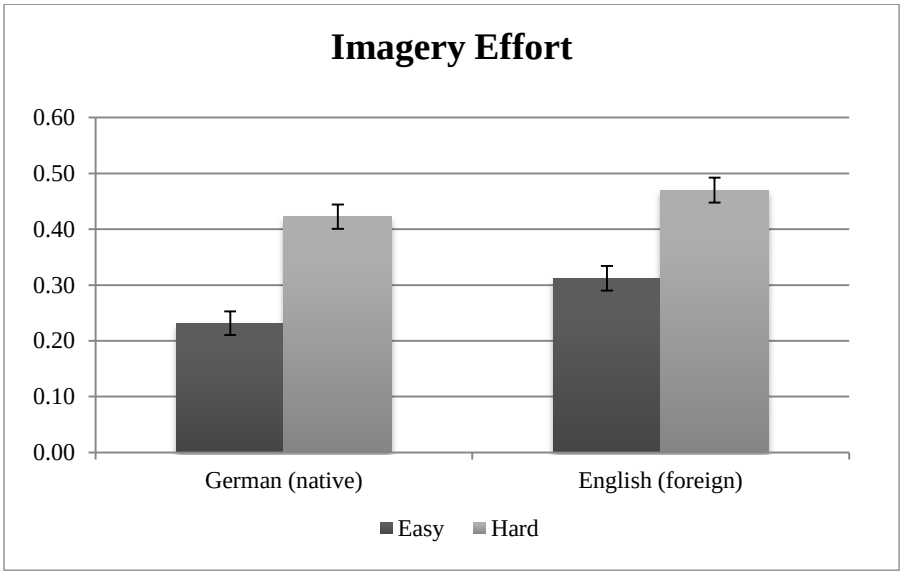

Figure 19: Effort imagining disease symptoms. The Y-axis represents effort ratings with higher values indicating greater effort imagining the symptoms. Error bars represent standard errors.

Discussion

The results from Studies 11 and 12 paint a somewhat complex picture of the relationship between language, imagery and perceptions of probability. We initially hypothesized that people using a foreign language should either rely less on mental imagery when making probability judgments, or else perceive all events to be less likely due to a reduction in the vividness of mental simulations. The former theory was confirmed in Study 11 such that participants were less affected by how easily symptoms could be imagined when using a foreign language. However, when directly asked how difficult it was to imagine the symptoms, those using a foreign language did not report any greater difficulty, contrary to the results from Studies 9 and 10. Study 12, on the other hand, showed more consistent imagery results such that those using a

[2] The same pattern of results was obtained for all measures when utilizing the more conservative screening criteria of including only participants with perfect translation scores.

foreign tongue rated their imagery to be more effortful, but did not replicate the previously observed Language x Difficulty interaction on likelihood judgments.

One potential explanation for this latter discrepancy may have something to do with the overall size of the imagery effect in the native language. While we observed a sizeable and significant effect of imagery difficulty in both experiments, the effect for native speakers was notably more pronounced in Study 11 relative to 12 (Ms = 24% difference vs. 14%). In both experiments, we observed about a 12% difference between Hard and Easy conditions when using the foreign language. It may thus be the case that we only observe a foreign language reduction in the size of the imagery effect when it is sufficiently pronounced in the native tongue.

In addition to the likelihood ratings, there is the curious inconsistency across experiments in terms of the effect of language on imagery ratings. The only notable difference between the two experiments was the wording of the question (how difficult vs. how effortful). It may be the case that such trivial wording differences could have affected the results, suggesting that future research may benefit from utilizing a more objective measure of mental imagery (such as in Study 10) coupled with ratings of likelihood in order to attempt a mediation analysis. Another possibility is that the effect of language on imagery depends on other factors such as foreign language proficiency or age of acquisition. Unfortunately, this data was only collected for participants in Study 11 and as such, cannot be directly addressed with the current data set. However, within Study 11, we do observe a marginal positive correlation between age of acquisition and the difficulty of imagery (r=0.13, p=.076) for those using their foreign language. Additionally, there was a small but marginal negative correlation between proficiency and imagery difficulty (r=-0.03, p=.09) for those using the foreign language, whereas no relationships were found for those using their native tongue. These results may suggest that

imagery may be less effortful for those who acquired the foreign language earlier as well as those who are more proficient. While there is no way to confirm this, it may be the case that the native German speakers in Study 12 were less proficient in English than the native Korean speakers in Study 11, thereby contributing to the discrepancy in language effects between the two studies. While further experimentation is necessary to definitively establish the effect, the present experiments provide tentative support for the theory that using a foreign language may reduce the reliance on mental imagery when making judgments of probability.

GENERAL DISCUSSION

"If you talk to a man in a language he understands, that goes to his head. If you talk to him in his language, that goes to his heart." –Nelson Mandela

One consequence of living in a global society is that there is a rich diversity of language backgrounds. Some of us live our entire lives with minimal exposure to other languages, while others, like a certain polyglottal professor I know, can converse in Hebrew, English, Arabic and a little Italian. Most of us lie somewhere in between, with various levels of formal education, proficiency, and cultural and emotional associations. This heterogeneity can make multilingual research challenging as a multitude of factors can affect an individual's relationship to their languages and how they subsequently affect behavior. Yet despite all of the noise and variance, an emerging pattern demonstrates that our native tongue is unlike every other language we may acquire.

Studies 1-6 demonstrated that in the moral domain, using a foreign language does not increase utilitarian cost-benefit considerations, but rather decreases adherence to deontological rules. These studies were among the first to tease these two processes apart and suggest that foreign language effects may emerge as a result of engaging less emotion, rather than more deliberation. To the extent that our findings generalize beyond moral dilemmas, they generate further predictions about the boundary conditions of foreign language use on decision-making. For example, Stanovich and West (2008) distinguished between decision-making biases that are or are not correlated with cognitive ability. Biases correlated with cognitive ability such as hindsight bias, outcome bias, belief bias for syllogistic reasoning, are likely reflective of deliberative System-2 processing, while those not correlated with cognitive ability such as sunk cost effects, conjunction fallacies, anchoring and adjustment may be reflective of heuristic

System-1 processing. An interesting question for future research is whether foreign language use attenuates decision biases that are associated with System-1 but not others associated with System-2.

Despite indications that using a foreign language decreases what are often thought of as emotional considerations, Studies 7 and 8 failed to detect differences in physiological arousal and parasympathetic activity. One interpretation of this null effect may be along the lines of those proposed by Lazar, Stern & Cohen (2014), who suggest that the added stress of using a foreign language may essentially cancel out the reduced emotional resonance of the task itself. Another possibility is that using a foreign language reduces reliance on heuristic System-1 processes without dramatically influencing emotional arousal. Rather than a strong emotional response, what may be altered is *affect*, or what Slovic describes as a "faint whisper of emotion… demarcating a positive or negative quality of a stimulus (Slovic et al., 2004). To use the authors' example, reading words such as "treasure" or "hate" will rapidly elicit a sense of goodness or badness, but may not be accompanied by a full-blown emotional experience. It may thus be the case that using a foreign language elicits these affective reactions either less automatically or less clearly, resulting in different evaluations and choices without notable changes in physiological activity.

Such a mechanism could be consistent with the results of Study 8 for which we observe an overall increase in both willingness to pay and willingness to accept in a foreign language, while no differences were found in physiological activity. Using a similar endowment paradigm, Lerner, Small & Loewenstein (2004) found that inducing a sad mood reduced willingness to accept and increased willingness to pay. While the authors did not take measures of physiological arousal, past research suggests that sadness is a relatively low-arousal emotion that

produces less dramatic changes in physiological activity compared to emotions such as anger (e.g. Ekman, Levenson, & Friesen, 1983; Shields, 1984). As such, it appears that changes in even low-arousal affects can change buying and selling decisions. This suggests that using a foreign language may influence choice by altering affect, even in the absence of physiological changes.

While a full-blown emotional experience may not be necessary to elicit a foreign language effect, it does appear that the most robust effects emerge when dealing with more obviously emotional topics such as moral dilemmas. As such, it may be the case that physiological differences could be observed when dealing with more emotionally evocative stimuli. Additionally, future research may benefit from utilizing a paradigm that does not require foreign language production (as in Caldwell-Harris & Ayçiçeği-Dinn, 2009 and Duñabeitia & Costa, 2015) to minimize performance anxiety that could counteract the effect.

So what makes certain stimuli more emotionally evocative than others? As discussed in Study 9, one factor appears to be how vividly we can imagine a given scene or scenario (e.g. Bartels, 2008). Studies 9, 10 and 12 demonstrate that this imagery is reduced in a foreign language, and in some cases, can partially account for observed changes in decision-making. This finding also provides a mechanistic explanation for why using a foreign tongue is less emotional than a native language, and may help guide future investigations into where and when language effects are likely to emerge. For instance, we may expect the strongest language effects for situations that are normally imagined quite vividly, such as near future events relative to far future events (D'Argembeau, & Van der Linden, 2004). Additionally, it raises the possibility that simply being a foreign language context would not be sufficient, but rather that the content of the choice should be processed in the foreign tongue for effects to emerge. Lastly, future research should systematically investigate whether or not there are comparable foreign language

reductions in vividness across all modalities. Recent work by our lab suggests that there are not. Using a self-report paradigm, we find that people using a foreign language experienced less vivid simulations for visual, auditory, tactile and motor sensations, but not for gustatory or olfactory (Hayakawa & Keysar, *under review*). This may in part explain why we found inconsistent effects of language for Studies 11 and 12 in which we asked participants to simulate symptoms of a disease rather than a visual scene. By investigating the types of stimuli that are especially difficult to simulate in a foreign language, we may gain a better understanding of the processes responsible for foreign language effects specifically, as well as the relationship between language and imagery more generally.

As noted in the introduction to Study 9, one reason why using a foreign language reduces mental imagery may have to do with language-dependent memory (Marian & Neisser, 2000), a form of encoding specificity (Thomson & Tulving, 1970). Imagery may not come as easily or vividly if the images in memory were encoded in a different language than the one being used to retrieve them. Such a mechanism could conceivably extend beyond mental imagery and apply to other learned concepts such as norms, values, and stereotypes. This could provide an alternative or additional explanation for why using a foreign language reduces adherence to deontological rules. While Study 9 provides evidence that reduced mental imagery partially accounts for this phenomenon, using a foreign language may additionally fail to activate the very concept of the deontological rule as clearly as in a native tongue. This would be consistent with Geipel et al. (2015b) and Pan & Patel's (2016) suggestion that using a foreign language primarily affects behavior by reducing the salience of norms. As such, a challenge of future research will be to partial out the extent to which using a foreign language alters how we react to norms and their violations versus the prevention of norms from becoming activated in the first place.

The fact that people's choices change depending on language is consistent with the idea that preferences are constructed during elicitation rather than inherent predilections that are uncovered (Slovic, 1995). As described by Gregory, Lichtenstein and Slovic (1993), it may be more appropriate to think of decision makers as architects who are building their preferences rather than archeologists who are unearthing them. Our research suggests that language provides key materials for building such preferences. Even robust decision biases can be attenuated or eliminated when using a foreign language, suggesting that much of what we have learned about decision-making may be contingent on the use of the native tongue. These findings thus demonstrate the intricate and reciprocal relationships that exist between language, cognition and choice.

In 2010, over 25% of physicians and surgeons were born outside of the United States (MPI analysis of American Community Survey, 2010). In the same year, there were approximately 40 million foreign-born residents in the United States, 44% of whom were naturalized citizens, making them eligible to serve on juries (U.S. Census Bureau, 2010). In 2006, the European Commission estimated that 56% of individuals in Europe were able to speak at least one language other than their native tongue (European Commission Special Eurobarometer, 2006). These numbers have surely increased in subsequent years and demonstrate the relevance of researching the effects of using a foreign language. In addition to continuing investigations into the underlying mechanisms, it will be important to extend laboratory research into the field and to more high-stakes contexts to examine how language impacts choices in domains such as medical and legal decision making. In some cases, the emotional distance afforded by a foreign language may help individuals make a more reasoned choice. In other cases, the strong intuitions provided by a native tongue may prove to be

advantageous. While the usefulness of a particular language will depend greatly on the context, it is important to keep in mind that acquiring a second language can do more than help you show off during your trip to Barcelona. As expressed by the Italian director Federico Fellini, “a different language is a different vision of life.”

REFERENCES

Addis, D. R., Wong, A. T., & Schacter, D. L. (2007). Remembering the past and imagining the future: Common and distinct neural substrates during event construction and elaboration. *Neuropsychologia*, *45*(7), 1363-1377.

Aleman, A. (2005). Feelings you can't imagine: towards a cognitive neuroscience of alexithymia. *Trends in cognitive sciences*, *9*(12), 553-555.

Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General*, *136*(4), 569-576.

Amit, E., & Greene, J. D. (2012). You see, the ends don't justify the means visual imagery and moral judgment. *Psychological science*, *23*(8), 861-868.

Appelhans, B. M., & Luecken, L. J. (2006). Heart rate variability as an index of regulated emotional responding. *Review of general psychology*, *10*(3), 229-240.

Ashcraft, M. H., & Krause, J. A. (2007). Working memory, math performance, and math anxiety. *Psychonomic Bulletin & Review*, *14*(2), 243–248.

Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology: General*, *129*(1), 126-145.

Barratt, E. S., & Patton, J. H. (1983). Impulsivity: Cognitive, behavioral, and psychophysiological correlates. *Biological Bases of Sensation Seeking, Impulsivity, and Anxiety*, 77-116.

Bartels, D. M. (2008). Principled moral sentiment and the flexibility of moral judgment and decision making. *Cognition*, *108*, 381–417.

Benet-Martínez, V., Leu, J., Lee, F., & Morris, M. W. (2002). Negotiating biculturalism cultural frame switching in biculturals with oppositional versus compatible cultural identities. *Journal of Cross-Cultural Psychology*, *33*(5), 492-516.

Benjamin, D., & Shapiro, J. (2005). Does cognitive ability reduce psychological bias?. *JEL manuscript*: J24, D14, C91.

Berlin, H. (2004). Impulsivity, time perception, emotion and reinforcement sensitivity in patients with orbitofrontal cortex lesions. *Brain*, *127*(5), 1108-1126.

Berntson, G. G., Thomas Bigger, J., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., Nagaraja, H. N., Porges, S.W., Saul, J. P., Stone, P.H., & Van Der Molen, M.W. (1997). Heart rate variability: origins, methods, and interpretive caveats. *Psychophysiology*, *34*(6), 623-648.

Bhatt, R., Williams, D. P., Jarczok, M. N., Koenig, J., & Thayer, J. F. (2015). Heart Rate Variability Predicts the Outcome of Risk Taking Behavior in Judgment and Decision Making. In Annual Meeting of the American Psychosomatic Society, Savannah, Georgia.

Bialystok, E., Craik, F. I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and aging*, *19*(2), 290-303.

Boroditsky, L. (2001). Does language shape thought?: Mandarin and English speakers' conceptions of time. *Cognitive Psychology*, *43*(1), 1-22.

Brislin, R. W. (1970). Back–Translation for Cross–Cultural Research. *Journal of Cross–Cultural Psychology, 1*(3), 185–216.

Butler, E. A., Wilhelm, F. H., & Gross, J. J. (2006). Respiratory sinus arrhythmia, emotion, and emotion regulation during social interaction. *Psychophysiology*, *43*(6), 612-622.

Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, *125*(3), 367-383.

Caldwell-Harris, C. L., & Ayçiçeği-Dinn, A. (2009). Emotion and lying in a non-native language. *International Journal of Psychophysiology, 71*(3), 193-204.

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.

Carroll, J. S. (1978). The effect of imagining an event on expectations for the event: An interpretation in terms of the availability heuristic. *Journal of experimental social psychology*, *14*(1), 88-96.

Caruso, E. M. (2008). Use of experienced retrieval ease in self and social judgments. *Journal of Experimental Social Psychology*, *44*(1), 148-155.

Casasanto, D. (2008). Who's afraid of the big bad Whorf? Crosslinguistic differences in temporal language and thought. *Language Learning*, *58*(s1), 63-79.

Čavar, F., & Tytus, A. E. (2017). Moral judgement and foreign language effect: when the foreign language becomes the second language. *Journal of Multilingual and Multicultural Development*, 1-12.

Chan, Y. L., Gu, X., Ng, J. C. K., & Tse, C. S. (2016). Effects of dilemma type, language, and emotion arousal on utilitarian vs deontological choice to moral dilemmas in Chinese–English bilinguals. *Asian Journal of Social Psychology*.

Cipolletti, H., McFarlane, S., & Weissglass, C. (2016). The moral foreign-language effect. *Philosophical Psychology, 29*(1), 23-40.

Conway, P., & Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: A process dissociation approach. *Journal of Personality and Social Psychology, 104*(2), 216–235.

Conway, P. & Rosas, A. (in preparation). Distinguishing between Harm Avoidance and Outcome Maximization in a New Battery of Process Dissociation Dilemmas, Accounting for Guilty Victims, Fated Victims, and Selfishness.

Costa, A., Foucart, A., Arnon, I., Aparici, M., & Apesteguia, J. (2014a). "Piensa" twice: on the foreign language effect in decision making. *Cognition, 130*(2), 236-254.

Costa, A., Foucart, A., Hayakawa, S., Aparici, M., Apesteguia, J., Heafner, J., & Keysar, B. (2014b). Your morals depend on language. *PloS one, 9*(4), e94842.

Costa, A., Vives, M. L., & Corey, J. D. (2017). On language processing shaping decision making. *Current Directions in Psychological Science, 26*(2), 146-151.

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.* Advance online publication. http://dx.doi.org/10.1037/xlm0000356

Cushman, F. (2013). Action, outcome, and value a dual–system framework for morality. *Personality and Social Psychology Review, 17*(3), 273–292.

D'Argembeau, A., & Van der Linden, M. (2004). Phenomenal characteristics associated with projecting oneself back into the past and forward into the future: Influence of valence and temporal distance. *Consciousness and cognition, 13*(4), 844-858.

Dadds, M., Hawes, D., Schaefer, B., & Vaka, K. (2004). Individual differences in imagery and reports of aversions. *Memory, 12*(4), 462-466.

Damasio, A. R. (1994). Descartes' error. New York: Putnam

De Brigard, F., Szpunar, K. K., & Schacter, D. L. (2013). Coming to grips with the past: Effect of repeated simulation on the perceived plausibility of episodic counterfactual thoughts. *Psychological Science, 24*(7), 1329-1334.

De Langhe, B, Puntoni, S, Fernandes, D & Van Osselaer, S. (2011). The anchor contraction effect in international marketing research. *Journal of Marketing Research,* 48(2), 366-380.

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science, 313*(5787), 684-687.

Dewaele, J. M. (2004). The emotional force of swearwords and taboo words in the speech of multilinguals. *Journal of multilingual and multicultural development, 25*(2-3), 204-222.

Dewaele, J. M. (2008). The emotional weight of I love you in multilinguals' languages. *Journal of Pragmatics, 40*(10), 1753-1780.

DeWall, C. N., Chester, D. S., & White, D. S. (2015). Can acetaminophen reduce the pain of decision-making?. *Journal of Experimental Social Psychology, 56,* 117-120.

Díaz-Lago, M. & Matute, H. (2014). Foreign language reduces the illusion of causality. Paper Presented at the Joint Meeting of the Sociedad Española de Psicología Experimental (SEPEX) and Sociedad Española de Psicofisiología y Neurociencia Cognitiva y Afectiva (SEPNECA). Murcia, Spain.

Dickert, S., Sagara, N., & Slovic, P. (2011). Affective motivations to help others: A two-stage model of donation decisions. *Journal of Behavioral Decision Making, 24*(4), 361-376.

Dolscheid, S., Shayan, S., Majid, A., & Casasanto, D. (2013). The thickness of musical pitch: Psychophysical evidence for linguistic relativity. *Psychological Science, 24*(5), 613-621.

Duñabeitia, J. A., & Costa, A. (2015). Lying in a native and foreign language. *Psychonomic Bulletin & Review, 22*(4), 1124-1129.

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science, 221*(4616), 1208-1210.

European Commission Special Eurobarometer. (2006). Europeans and their languages. Retrieved from http://ec.europa.eu/public_opinion/archives/ebs/ebs_243_en.pdf

Favreau, M., & Segalowitz, N. S. (1983). Automatic and controlled processes in the first-and second-language reading of fluent bilinguals. *Memory & Cognition, 11*(6), 565-574.

Foot, P. (1978). The problem of abortion and the doctrine of the double effect in virtues and vices. Oxford: Basil Blackwell.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25–42.

Frey, D. P., & Gamond, L. (2015). Second language feedback reduces the hot hand fallacy, but why?. *Journal of Neuroscience, 35*(34), 11766-11768.

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.

Gao, S., Zika, O., Rogers, R. D., & Thierry, G. (2015). Second language feedback abolishes the "hot hand" effect during even-probability gambling. *Journal of Neuroscience, 35*(15), 5983-5989.

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.

Geipel, J., Hadjichristidis, C., & Surian, L. (2015b). How foreign language shapes moral judgment. *Journal of Experimental Social Psychology, 59*, 8-17.

Greene, J., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science 293*, 2105–2108.

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.

Gregory, W. L., Cialdini, R. B., & Carpenter, K. M. (1982). Self-relevant scenarios as mediators of likelihood estimates and compliance: Does imagining make it so?. *Journal of personality and social psychology*, *43*(1), 89-99.

Gregory, R., Lichtenstein, S., & Slovic, P. (1993). Valuing environmental resources: a constructive approach. Journal of Risk and Uncertainty, 7(2), 177-197.

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.

Hadjichristidis, C., Geipel, J., & Surian, L. (2016). How foreign language affects decisions: Rethinking the brain-drain model. *Journal of International Business Studies*, 1-7.

Haigh, M. S., & List, J. A. (2005). Do professional traders exhibit myopic loss aversion? An experimental analysis. *The Journal of Finance*, *60*(1), 523-534.

Hansen, A. L., Johnsen, B. H., & Thayer, J. F. (2003). Vagal influence on working memory and attention. *International Journal of Psychophysiology*, *48*(3), 263-274.

Harris, C. L., Aycicegi, A., & Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. *Applied Psycholinguistics, 24*(04), 561-579.

Hayakawa, S., Costa, A., Foucart, A., & Keysar, B. (2016). Using a foreign language changes our choices. *Trends in Cognitive Sciences, 20*(11), 791-793.

Hayakawa, S. & Keysar, B. (under review). Morality and mental imagery in a foreign language.

Hassabis, D., Kumaran, D., Vann, S. D., & Maguire, E. A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. *Proceedings of the National Academy of Sciences*, 104(5), 1726-1731.

Hernandez, I., & Preston, J. L. (2013). Disfluency disrupts the confirmation bias. *Journal of Experimental Social Psychology*, *49*(1), 178-182.

Holmes, E. A., & Mathews, A. (2005). Mental imagery and emotion: a special relationship?. *Emotion*, *5*(4), 489-497.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, *92*(5), 1644-1655.

Hsu, C. T., Jacobs, A. M., & Conrad, M. (2015). Can Harry Potter still put a spell on us in a second language? An fMRI study on reading emotion-laden literature in late bilinguals. *Cortex*, *63*, 282-295.

Im Shin, H., & Kim, J. (2017). Foreign Language Effect and Psychological Distance. *Journal of Psycholinguistic Research*, 1-14.

Isaac, A., Marks, D. F., & Russell, D. G. (1986). An instrument for assessing imagery of movement: The Vividness of Movement Imagery Questionnaire (VMIQ). *Journal of Mental Imagery*.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*(5), 513–541.

Johnson, E. J., Häubl, G., & Keinan, A. (2007). Aspects of endowment: a query theory of value construction. *Journal of experimental psychology: Learning, memory, and cognition*, *33*(3), 461.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological review, 80*(4), 237-251.

Kahneman, D. & Tversky A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica, 47*, 263–292.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *The journal of economic perspectives*, *5*(1), 193-206.

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

Kassam, K. S., Koslov, K., & Mendes, W. B. (2009). Decisions under distress: Stress profiles influence anchoring and adjustment. *Psychological Science, 20*(11), 1394-1399.

Keysar, B., Hayakawa, S. L., & An, S. G. (2012). The foreign-language effect thinking in a foreign tongue reduces decision biases. *Psychological science*, *23*(6), 661-668.

Klesse, A. K., Levav, J., & Goukens, C. (2015). The effect of preference expression modality on self-control. *Journal of Consumer Research*, *42*(4), 535-550.

Knutson, B., Wimmer, G. E., Rick, S., Hollon, N. G., Prelec, D., & Loewenstein, G. (2008). Neural antecedents of the endowment effect. *Neuron*, *58*(5), 814-822.

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., & Damasio, A. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature*, *446*(7138), 908-911.

Kogut, T., & Ritov, I. (2005). The "identified victim" effect: An identified group, or just a single individual?. *Journal of Behavioral Decision Making*, *18*(3), 157-167.

Lazar, J. N., Stern, A., & Cohen, R. (2014). Decision-making in foreign language reduces emotional arousal. *Psychology, 5*(19), 2180-2188.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., & Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied, 8*(2), 75-84.

Lequerica, A, Rapport, L, Axelrod, B, Telmet, K & Whitman, R. (2002). Subjective and objective assessment methods of mental imagery control: Construct validations of self-report measures. *Journal of Clinical and Experimental Neuropsychology* 24(8), 1103-1116.

Lerner, J. S., Small, D. A., & Loewenstein, G. (2004). Heart strings and purse strings carryover effects of emotions on economic decisions. *Psychological science, 15*(5), 337-341.

Levi, A. S., & Pryor, J. B. (1987). Use of the availability heuristic in probability estimates of future events: The effects of imagining outcomes versus imagining reasons. *Organizational Behavior and Human Decision Processes, 40*(2), 219-234.

Lipsey, M.W. & Wilson, D. B. (2001). Practical meta-analysis (Vol. 49). Thousand Oaks, CA: Sage publications.

Lucy, J. A. (2004). Language, culture, and mind in comparative perspective. *Language, culture and mind*, 1-22.

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate behavioral research*, 39(1), 99-128.

Marian, V., & Neisser, U. (2000). Language-dependent recall of autobiographical memories. *Journal of Experimental Psychology: General, 129*, 361–368.

Mathias, C. W., & Stanford, M. S. (2003). Impulsiveness and arousal: heart rate under conditions of rest and challenge in healthy males. *Personality and Individual Differences, 35*(2), 355-371.

Matsumoto, A., & Stanny, C. (2006). Language-dependent access to autobiographical memory in Japanese-English bilinguals and US monolinguals. *Memory*, 14(3), 378-390.

Mehta, Z., Newcombe, F., & De Haan, E. (1992). Selective loss of imagery in a case of visual agnosia. *Neuropsychologia, 30*(7), 645-655.

Miltner, W. H., Braun, C. H., & Coles, M. G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a "generic" neural system for error detection. *Journal of cognitive neuroscience, 9*(6), 788-798.

MPI analysis of American Community Survey (2010). *Foreign-Born Health Care Workers in the United States.* Retrieved from http://www.migrationpolicy.org/article/foreign-born-health-care-workers-united-states#4

Nakamura, K. (2015). Effect is sure, but explanation is unsure: Closer investigation of the foreign language effect with Japanese participants. In CogSci. https://mindmodeling.org/cogsci2015/papers/0295/paper0295.pdf

Nakamura, K. (2016). Our morals really depends on our language: The foreign language effect within participants. https://mindmodeling.org/cogsci2016/papers/0175/paper0175.pdf

Oganian, Y., Korn, C. W., & Heekeren, H. R. (2016). Language switching—but not foreign language use per se—reduces the framing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42(*1), 140-147.

Okuda, J., Fujii, T., Ohtake, H., Tsukiura, T., Tanji, K., Suzuki, K., ... & Yamadori, A. (2003). Thinking of the future and past: The roles of the frontal pole and the medial temporal lobes. *Neuroimage*, *19*(4), 1369-1380.

Orwell, George. 1984. New York: Penguin, 1949.

Pan, P., & Patel, C. (2016). The influence of native versus foreign language on Chinese subjects' aggressive financial reporting judgments. *Journal of Business Ethics*, 1-16.

Paulus, M. P., & Stein, M. B. (2006). An insular view of anxiety. *Biological psychiatry, 60*(4), 383-387.

Pavlenko, A. (2005). Emotions and multilingualism. New York, NY: Cambridge University Press.

Pavlenko, A. (2012). Affective processing in bilingual speakers: Disembodied cognition? *International Journal of Psychology, 47*, 405–428.

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (2003). Cognitive load in reading a foreign language text with multimedia aids and the influence of verbal and spatial abilities. *Computers in Human Behavior, 19*(2), 221–243.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior research methods, instruments, & computers*, 36(4), 717-731.

Prelec, D., & Loewenstein, G. (1998). The red and the black: Mental accounting of savings and debt. *Marketing science*, *17*(1), 4-28.

Puntoni, S., De Langhe, B., & Van Osselaer, S. M. (2009). Bilingualism and the emotional intensity of advertising language. *Journal of Consumer Research, 35*(6), 1012-1025.

Rabin, M., & Thaler, R. H. (2001). Anomalies: risk aversion. *Journal of Economic perspectives*, 219-232.

Ramírez-Esparza, N., Gosling, S. D., Benet-Martínez, V., Potter, J. P., & Pennebaker, J. W. (2006). Do bilinguals have two personalities? A special case of cultural frame switching. *Journal of research in personality*, *40*(2), 99-120.

Rand, D. G., Greene, J. D., & Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature, 489,* 427–430.

Redelmeier, D. A., & Tversky, A. (1990). Discrepancy between medical decisions for individual patients and for groups. *Preference, Belief, and Similarity*, 887-893.

Robinson, P. (2001). Individual differences, cognitive abilities, aptitude complexes and learning conditions in second language acquisition. *Second Language Research 17*(4), 368–392.

Samanez-Larkin, G. R., Hollon, N. G., Carstensen, L. L., & Knutson, B. (2008). Individual differences in insular sensitivity during loss anticipation predict avoidance learning. *Psychological Science*, *19*(4), 320-323.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science, 300*(5626), 1755-1758.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. *Nature Reviews Neuroscience*, 8(9), 657-661.

Schrauf, R. W., & Rubin, D. C. (1998). Bilingual autobiographical memory in older adult immigrants: A test of cognitive explanations of the reminiscence bump and the linguistic encoding of memories. *Journal of Memory and Language*, 39(3), 437-457.

Schwarz, N. (2011). Feelings-as-information theory. *Handbook of theories of social psychology*, *1*, 289-308.

Sherman, S. J., Cialdini, R. B., Schwartzman, D. F., & Reynolds, K. D. (1985). Imagining can heighten or lower the perceived likelihood of contracting a disease the mediating effect of ease of imagery. *Personality and Social Psychology Bulletin*, *11*(1), 118-127.

Shields, S. A. (1984). Reports of bodily change in anxiety, sadness, and anger. *Motivation and Emotion, 8(*1), 1-21.

Simonson, I., & Drolet, A. (2004). Anchoring effects on consumers' willingness-to-pay and willingness-to-accept. *Journal of consumer research, 31*(3), 681-690.

Shiv, B., & Huber, J. (2000). The impact of anticipating satisfaction on consumer choice. *Journal of Consumer Research, 27*(2), 202-216.

Shiv, B., Lowenstein, G., Bechara, A., Damasio, H., & Damasio, A. (2005) Investment behavior and the negative side of emotion. *Psychological Science, 16*, 435-439.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*, 3–22.

Slovic, P. (1995). The construction of preference. *American psychologist, 50*(5), 364-371.

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk analysis, 24*(2), 311-322.

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

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23*, 645–665.

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of personality and social psychology, 94*(4), 672-695.

Thomson, J. (1985). The trolley problem. *Yale Law Journal, 94*, 1395–1415.

Sütterlin, S., Herbert, C., Schmitt, M., Kübler, A., & Vögele, C. (2011). Frames, decisions, and cardiac–autonomic control. *Social Neuroscience*, *6*(2), 169-177.

Szpunar, K. K., Watson, J. M., & McDermott, K. B. (2007). Neural substrates of envisioning the future. *Proceedings of the National Academy of Sciences*, *104*(2), 642-647.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization, 1*(1), 39-60.

Thayer, J. F., & Lane, R. D. (2009). Claude Bernard and the heart–brain connection: Further elaboration of a model of neurovisceral integration. *Neuroscience and Biobehavioral Reviews, 31*, 81–88.

Thayer, J. F., Åhs, F., Fredrikson, M., Sollers, J. J., & Wager, T. D. (2012). A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. *Neuroscience & Biobehavioral Reviews*, *36*(2), 747-756.

Thompson, J. J. (1985). Double effect, triple effect and the trolley problem: Squaring the circle in looping cases. *Yale Law Journal*, *94*(6), 1395-1415.

Thomson, D. M., & Tulving, E. (1970). Associative encoding and retrieval: Weak and strong cues. *Journal of Experimental Psychology, 86*(2), 255-262.

Tong, L. C., Ye, K. J., Asai, K., Ertac, S., List, J. A., Nusbaum, H. C., & Hortaçsu, A. (2016). Trading experience modulates anterior insula to reduce the endowment effect. *Proceedings of the National Academy of Sciences*, *113*(33), 9238-9243.

Traczyk, J., Sobkow, A., & Zaleskiewicz, T. (2015). Affect-laden imagery and risk taking: the mediating role of stress and risk perception. *PloS one*, *10*(3), e0122226.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie canadienne*, 26(1), 1-12.

Turula, A. (2016). Thinking in a foreign language, fast and slow. *Polish Psychological Bulletin, 47*(2), 228-232.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology, 5*(2), 207-232.

Tversky, A., & Kahneman, D. (1975). Judgment under uncertainty: Heuristics and biases. In Utility, probability, and human decision making (pp. 141-162). Springer Netherlands.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453–458.

Urbig, D., Terjesen, S., Procher, V., Muehlfeld, K., & Van Witteloostuijn, A. (2016). Come on and take a free ride: Contributing to public goods in native and foreign language settings. *Academy of Management Learning & Education, 15*(2), 268-286.

U.S. Census Bureau (2010). Census of Population, 1850 to 2000, and the American Community Survey, 2010. Retrieved from https://www.census.gov/newsroom/pdf/cspan_fb_slides.pdf

Wang, Y., Highhouse, S., Lake, C. J., Petersen, N. L., & Rada, T. B. (2015). Meta-analytic investigations of the relation between intuition and analysis. *Journal of Behavioral Decision Making.*

Weaver, R., & Frederick, S. (2012). A reference price theory of the endowment effect. *Journal of Marketing Research, 49*(5), 696-707.

Weber, B., Aholt, A., Neuhaus, C., Trautner, P., Elger, C. E., & Teichert, T. (2007). Neural evidence for reference-dependence in real-market-transactions. *Neuroimage*, *35*(1), 441-447.

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences, 104*(19), 7780-7785.

Winskel, H., Ratitamkul, T., Brambley, V., Nagarachinda, T., & Tiencharoen, S. (2016). Decision-making and the framing effect in a foreign and native language. *Journal of Cognitive Psychology, 28*(4), 427-436.

Wu, Y. J., & Thierry, G. (2012). How reading in a second language protects your heart. The *Journal of Neuroscience, 32*, 6485–6489.

Wu, Y., & Zhou, X. (2009). The P300 and reward valence, magnitude, and expectancy in outcome evaluation. *Brain research, 1286*, 114-122.

APPENDIX

STUDIES 1-6

Proficiency Effects.

We conducted a series of analyses to examine the possible effect of foreign language proficiency on the *U* and *D* parameters[1]. Lower levels of proficiency in a foreign language could increase cognitive load (Plass, Chun, Mayer, & Leutner, 2003). This in turn, could affect moral judgment, particularly the *U* parameter as utilitarian responding is thought to be the more cognitively expensive and may thus be reduced when resources are taxed (Greene et. al, 2008). Such an explanation could potentially explain why in some of our experiments participants using the foreign language were significantly less utilitarian than those using their native tongue. To test this hypothesis, we first calculated a "Relative Proficiency" score by subtracting the average reported fluency rating for the foreign language from the native language. A high score thus indicates that the foreign language proficiency was notably lower than that of the native tongue. We then separately regressed *U* and *D* scores onto experimental condition (0 = native language, 1 = foreign language), relative proficiency, and the interaction between the two. Consistent with the cognitive load hypothesis, we find that there is a significant or nearly significant effect of proficiency on *U* scores for two of out the three experiments for which we observed a foreign language reduction in utilitarianism (see Table A1). Additionally, we find significant proficiency x language interactions for those studies as the relationship between proficiency and *U* only holds for those using the foreign language during the experiment. We find no relationship between proficiency and the *D* parameter, which is what would be expected if deontological responding relies on relatively automatic System-I processes.

[1] Possible gender effects were also examined, but no reliable effects emerged.

	Proficiency		Proficiency x Language	
	b (SE)	*p-value*	*b (SE)*	*p-value*
Study 1	.048 (.037)	0.198	-.020 (.024)	0.388
Study 2	.080 (.042)	0.059¶	-.058 (.027)	0.033*
Study 3	-.021 (.067)	0.749	-.007 (.041)	0.856
Study 4	-.005 (.047)	0.911	-.014 (.032)	0.663
Study 5	-.023 (.051)	0.646	.001 (.033)	0.967
Study 6	.120 (.059)	0.042*	-.091 (.038)	0.017*

Table A1. Proficiency on U. ¶ $p \leq .10$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

	Proficiency		Proficiency x Language	
	b (SE)	*p-value*	*b (SE)*	*p-value*
Study 1	-.008 (.037)	0.822	.007 (.023)	0.756
Study 2	.024 (.046)	0.609	-.026 (.029)	0.282
Study 3	.103 (.074)	0.166	-.054 (.045)	0.234
Study 4	-.042 (.046)	0.361	.011 (.031)	0.72
Study 5	-.078 (.055)	0.16	.037 (.036)	0.297
Study 6	.008 (.065)	0.901	-.027 (.042)	0.518

Table A2: Proficiency on D

STUDY 7

Gender Effects.

Given previous literature demonstrating that men are often more risk-seeking than women (see Byrnes, Miller & Schafer, 1999 for meta-analysis), we examined the effect of gender on risk taking[2]. Indeed, we did find a marginal effect of Gender, though it was in the opposite direction than anticipated such that women took more gambles than men (Ms = 67% vs 60%; $\chi^2(1, N=97) = 3.38$, *p=.066; b= 0.41, SE=0.22*). There was no interaction between Gender and Language ($\chi^2(1, N=97) = 0.339$, *p=.561; b= 0.26, SE=0.44*). There was, however a significant Gender x Valence interaction such that men were more sensitive to expected value

[2] Possible proficiency effects were also examined, but no reliable effects emerged.

than women, as reflected by the difference in proportions of Good and Bad bets accepted (Ms = 36% vs 16%; (χ^2(1, N=97) = 47.50, *p<.001; b= -0.98, SE=0.14*). Additionally, there was a significant Gender x Domain interaction such that women were more affected by the domain of the decision, taking significantly more gambles in the Loss domain than the Gain domain relative to men (Ms = 21% vs 8% difference, χ2(1, N=97) = 27.34, *p*<.001; *b*= -0.73, SE=0.14). The main effect of language and the interaction with valence did not change as a result of including gender as a variable. However, including Gender as variable revealed a significant Language x Domain interaction such that those using a foreign language took significantly more bets in the loss domain relative to gain compared to those using their native tongue (Ms = 18% vs 14% difference; χ2(1, N=97) = 5.17, *p*=.023; *b*= -0.31, SE=0.14).

STUDY 8

Offer Decisions.

We examined the effect of language and task on the proportion of offers accepted. As noted in the methods, the offers were generated based on the slider values so that for each object, participants were presented with one offer at the stated value, 15 offers below and 15 above. As such, if the stated value reflected true intentions, we should expect participants to accept exactly 16 out of 31 offers for each object (51.6%). To examine whether this was the case, we ran a series of one-sample t-tests comparing the proportion of offers accepted to 51.6%. The data reveals that overall, participants accepted significantly more offers to sell than would be expected (M=61.0%; t(127)=6.21, *p*<.001), suggesting that participants were in fact willing to accept less than they stated. This was the case for both native and foreign language users (Ms = 63.1% and 59.0%, respectively) and the two language groups did not significantly differ from

each other (F(1, 126)=2.13, p=.147). On the other hand, participants accepted significantly fewer offers to buy than would be expected (M=39.3%; t(127)=-6.59, p<.001), suggesting that they were unwilling to pay as much as they stated. This was the case for both native and foreign language users (Ms = 37.0% and 42.0%, respectively) and the two groups did not differ from each other (F(1, 126)=1.70, p=.195).

Inferred Values (Offers).

Given the apparent discrepancy between stated and revealed values, we explored the effects of language and task on the values that can be inferred through the offer decisions. Within the selling blocks, we took the minimum accepted offer as the inferred selling value. Within the buying blocks, we took the maximum accepted offer as the inferred buying value. Just as for the stated (slider) values, we find a main effect of language such that those using a foreign language both demanded more to sell, and were more willing to pay to buy (F(1, 126)=10.44, p=.002). Additionally, we observe a main effect of task such that participants valued objects more when selling than buying, replicating the usual endowment effect (F(1, 126)=77.33, p<.001). Once again, we did not find a task x language interaction suggesting that the two languages had comparable endowment effects (F(1, 126)<.001, p=.990). As in the case of the stated values, some items were valued more highly than others (F(3, 378)=10.56, p<.001), but this did not interact with language (F(3,378)=1.30, p=.273). Also in line with the stated values, there was a significant item x task interaction such that some items evoked a stronger endowment effect than others (F(3, 378)=6.51, p<.001).

While the basic effects appear to be the same when looking at stated versus inferred values, we did find that the endowment effect was significantly reduced for revealed values

relative to stated values, as revealed by a significant type (stated vs inferred) x task (selling vs buying) interaction ($F(1, 126)=35.29$, $p<.001$). This did not, however, interact with language ($F(1, 126)=.19$, $p=.663$). This reduction in the endowment effect for inferred values appears to be driven by people being willing to accept significantly less than originally stated to sell their items ($F(1, 126)= 34.84$, $p<.001$). No difference between inferred and stated values were found for buying decisions ($F(1,126)=1.0$, $p=.319$). Again, neither of these effects were qualified by language.

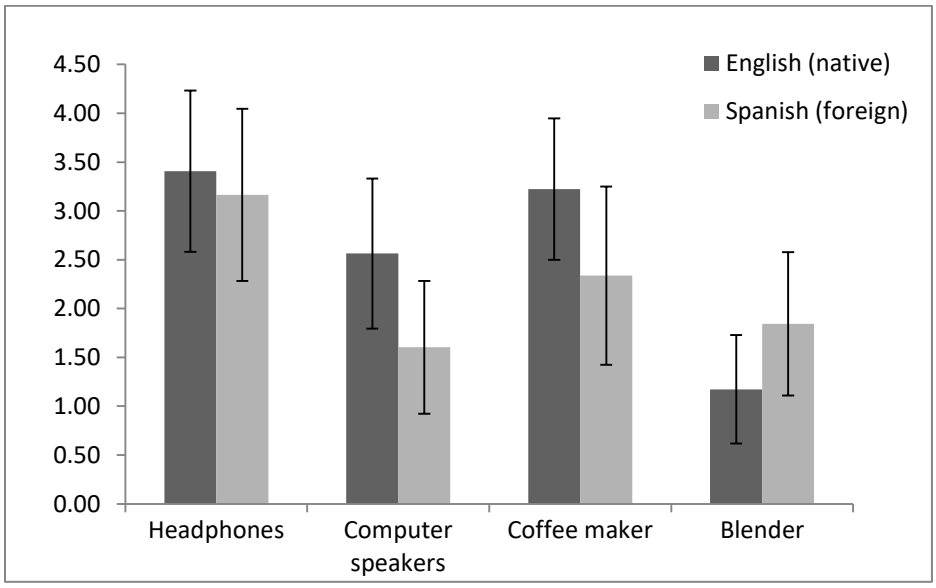

Figure A1: Difference in endowment effect between stated and inferred values .Values represent stated minus inferred values. Higher numbers indicate relatively greater endowment effect for stated (slider) values relative to inferred offer values.

Potential Moderators

Ownership.

We ran a series of analyses to examine whether the effect of language on the endowment effect was moderated by a number of variables. First we examined whether current ownership of

the item moderated the effect of language on the endowment effect. At the end of the experiment, participants were asked to state whether they currently owned each of the four items, and so a moderation analysis was conducted for each item separately. It may be the case the effect of language may only emerge when the decision is sufficiently consequential, such as when the object is not already owned. No moderating effect of ownership was found for headphones (b =2.19, SE=3.32, p=.509), the coffee maker (b =1.98, SE=3.31, p=.552), or the blender (b =-1.37, SE=3.31, p=.680). There was a significant moderating effect for speakers such that those using a foreign language had a significantly smaller endowment effect relative to native speakers for participants who currently owned speakers (N=54; M=2.92 vs 7.45, respectively; b =-4.53, SE=2.16, p=.038). No such effect was found for participants who did not (N= 70; Ms=9.70 vs 7.63; B=2.07, SE=1.89, p=.277). This latter result suggests that ownership status may play a role in predicting whether language will impact the endowment effect. However, given that this effect was found for only one of four objects, it should be interpreted with caution. No differences were found in likelihood of owning the objects across language groups (χ^2(1, N=124)=.55, p=.457).

Desirability.

In addition to indicating ownership, participants rated the desirability of each of the objects. It may be the case that language effects only emerge when the item is sufficiently desirable. However, no moderating effect of desirability was found for the headphones (b =.45, SE=.88, p=.609), coffee maker (b =-.52, SE=.90, p=.566), blender (b =3.16, SE=2.23, p=.158), or speakers (b =.32, SE=.94, p=.738). Additionally, there was no main effect of language in how desirable participants perceived the objects to be (F(1, 122)=1.05, p=.307). Desirability did significantly predict the size of the endowment effect such that the endowment effect became

increasingly smaller as desirability went up (b =-2.63, SE=.83, p=.002). However, this did not interact with language (b =.89, SE=1.19, p=.459). This relationship between desirability and the endowment effect appear to be driven by the buying condition such that, unsurprisingly, individuals are willing to pay more for items they desire (b =1.79, SE=.61, p=.004). Desirability did not, however, affect selling prices (b =-.83, SE=.79, p=.296). Once again, the effects of desirability did not interact with language for either buying (p=.569) or selling (p=.788).

Demographics.

We did not find moderation effects for either major (economics or not; b =-1.89, SE=3.99, p=.635) or gender (b =-2.15, SE=2.89, p=.459). While at first, it appeared that proficiency was a marginally significant moderator, with foreign language users less susceptible to the endowment effect only at higher levels of proficiency (b =-3.54, SE=1.64, p=.033), this effect disappeared after excluding one outlier (b =-2.53, SE=1.86, p=.176). All other effects remain the same after excluding this participant.

Heart rate and the endowment effect.

To examine whether heart rate predicted buying values, we separately regressed buying rating values onto resting heart rate and heart rate during the buying task. As can be seen in Table A3, neither measure was significantly predictive of buying values, either for explicitly stated slider values or inferred offer values. The same analysis was conducted for selling values and once again, neither measure was predictive. Lastly, the analysis was run again with the endowment effect (selling minus buying) as the dependent variable and resting heart rate and the difference between selling and buying heart rates as predictors. There was a marginal effect of endowment heart rate on the endowment effect such that the greater the difference in heart rate

between selling and buying blocks, the greater the difference in valuation, specifically for the stated slider values. This effect was not moderated by language (b =-.006, SE=.51, p=.99). No other effects were significant.

		Slider Values		Inferred Offer Values	
	Source	*b* (SE)	p-value	*b* (SE)	p-value
BUYING	Resting Heart Rate	-.09 (.06)	0.129	-.11 (.06)	0.099
	Buying Heart Rate	-0.08 (0.07)	0.273	-.06 (.08)	0.425
SELLING	Resting Heart Rate	.05 (.08)	0.535	.003 (.002)	0.101
	Selling Heart Rate	.09 (.09)	0.34	.003 (.002)	0.152
ENDOWMENT	Resting Heart Rate	.14 (.09)	0.111	.10 (.09)	0.28
	Endowment Heart Rate	.58 (.31)	0.067	.33 (.61)	0.59

Table A3: Effects of resting and average heart rate on buying price, selling price, and the endowment effect

RSA and the endowment effect.

The same set of analyses were conducted with RSA. As can be seen in Table A4, RSA at rest or during the task did not predict buying, selling or endowment values, either those which were explicitly stated or those which were inferred by the offers accepted.

		Slider Values		Inferred Offer Values	
	Source	*b* (SE)	p-value	*b* (SE)	p-value
BUYING	Resting RSA	.49 (.62)	0.207	.76 (.66)	0.256
	Buying RSA	.74 (.74)	0.321	.45 (.89)	0.616
SELLING	Resting RSA	.36 (.81)	0.658	.33 (.83)	0.691
	Selling RSA	-.59 (1.05)	0.34	-1.20 (.88)	0.184
ENDOWMENT	Resting RSA	-.13 (.92)	0.891	-.43 (.96)	0.656
	Endowment RSA	-.81 (1.34)	0.553	.52 (2.15)	0.811

Table A4*:* Effects of resting and average RSA on buying price, selling price, and the endowment effect

Moderation analyses revealed that resting RSA did not moderate the effect of language on the endowment effect for the stated values (b =.11, SE=.192, p=.953) or inferred values (b =.05, SE=1.99, p=.980).

STUDY 9

No effects of foreign language proficiency or gender were found for study 9.

STUDY 10

Proficiency Effects.

We examined the potential role of foreign language proficiency for performance overall (averaging across the tasks), as well as for each of the four tasks separately (Shape-Word, Category-Word, Shape-Picture, Category-Picture). To do so, we regressed each of these measures on to language condition, foreign language proficiency, and the interaction between the two.

Overall.

Collapsing across tasks, no main effect of foreign language proficiency emerged (b=-0.001, SE=.01, p=*.532*). However, there was a significant condition language x foreign proficiency interaction (b=-0.05, SE=.01, p<*.001*). This interaction reflects a pattern such that for those using the foreign language, those with higher foreign proficiency performed better overall (b=0.05, SE=.01, p<*.001*). Unsurprisingly, no relationship emerged between foreign proficiency and performance for those using their native language (b=-0.01, SE=.004, p=.246). As can be seen in Table A5, a similar pattern was found for each of the tasks separately, with the exception of Category-Picture, for which the interaction did not reach significance. It is of note

that in all cases, higher foreign language proficiency was associated with better performance when using the foreign language.

		Proficiency		Language x Proficiency	
	Task	*b* (SE)	p-value	*b* (SE)	p-value
Word	Shape	.01 (.01)	0.618	.07 (.02)	0.001
	Category	.01 (.01)	0.329	.05 (.02)	0.003
Picture	Shape	-.01 (.01)	0.345	.04 (.01)	0.004
	Category	.01 (.01)	0.405	.02 (.02)	0.187

Table A5: Effects of foreign language proficiency on performance

Gender Effects.

The same set of analyses were conducted with gender rather than foreign language proficiency. Overall, no effect of gender (p=.742) nor a gender x language interaction emerged (p=.207). The only notable gender effect was a nearly significant gender x language interaction for the Category-Word task (b=0.07, SE=.04, p=*.052*) such that women marginally outperformed men when using a foreign language (b=0.57, SE=.03, p=*.071*), but the native language (b=-0.01, SE=.02, p=*.460*). No other effects of gender were significant.

STUDIES 11 & 12

Study 11. No effects of either foreign language proficiency or gender were found for either measure of likelihood (disease/romantic relationship). See Chapter on Studies 11 and 12 for analysis of proficiency effects for imagery ratings. No gender effects were found for imagery ratings.

Study 12. No proficiency data was collected for participants in study 12. There was a significant main effect of gender on the perceived likelihood of contracting the disease such that women

thought it was more likely than men (F(1, 431)=8.37, p=.004, η^2=.02). This did not interact with either language (native vs. foreign; F(1,431)=.27, p=.607, η^2=.001), or target (self vs. other; F(1,431)=2.12, p=.146, η^2=.006). No main effect of gender emerged for imagery effortfulness (F(1,431)=2.04, p=.154, η^2=.005), however there was a significant language x gender interaction such that women found the imagery to be less effortful than men when using their native language, whereas no gender effects were found in the foreign language (F(1,431)=7.32, p=.007, η^2=.02).