Special Issue Article

QJEP

Quarterly Journal of Experimental Psychology
2019, Vol. 72(1) 60–75

DOI: 10.1177/1747021818774789
qjep.sagepub.com
SAGE

Low foreign language proficiency reduces optimism about the personal future

Yulia Oganian[1,2,3], Hauke R Heekeren[1,4] and Christoph W Korn[1,5]

Abstract

Optimistic estimates about the personal future constitute one of the best-described and most-debated decision biases related to emotion. Nevertheless, it has been difficult to isolate manipulations that reduce optimistic estimates. Eliciting estimates in a foreign language is a promising candidate manipulation because foreign language use alters decision biases in scenarios with emotional components. Consequently, we tested whether foreign language use reduces optimistic estimates. In a laboratory experiment, participants (n = 45) estimated their probability of experiencing life events either in their native language or a foreign language, in which they were highly proficient. We found no differences in these estimates or in the updating of these estimates after receiving feedback about the population baseline probability. Importantly, three online experiments with large sample sizes (ns = 706, 530, and 473) showed that using a foreign language with low proficiency reduced comparative optimism. Participants in the online experiments had diverse proficiency levels and were matched on a variety of control metrics. Fine-grained analyses indicated that low proficiency weakens the coupling between probability estimates and rated arousal. Overall, our findings suggest that an important decision bias can be reduced when using a foreign language with low proficiency.

Keywords

Decision-making biases; apparent irrationalities; foreign language effects; optimism; self

Received: 16 March 2017; revised: 9 April 2018; accepted: 11 April 2018

Ideally, human decisions should comply with rational and normative standards. Nevertheless, countless studies have indicated that humans consistently make decisions that violate normative standards (Kahneman, 2003). Prominent examples are apparent optimism and framing effects: Humans often hold (normatively unwarranted) optimistic expectations about their future (Sharot, 2011; Sharot & Garrett, 2016; Shepperd, Klein, Waters, & Weinstein, 2013) and succumb to the (normatively irrelevant) framing of choice situations (Kühberger, 1998; Tversky & Kahneman, 1981).

Mounting evidence suggests that framing effects are reduced in foreign language contexts, such as when participants read framing scenarios in a foreign language (Costa, Foucart, Arnon, et al., 2014; Keysar, Hayakawa, & An, 2012; Winskel, Ratitamkul, Brambley, Nagarachinda, & Tiencharoen, 2016) or when they unexpectedly switch between foreign and native languages (Oganian, Korn, & Heekeren, 2016). Furthermore, foreign language use alters gambling behaviour (Costa, Foucart, Arnon, Aparici, & Apesteguia, 2014; Gao, Zika, Rogers, & Thierry, 2015; Keysar et al., 2012), influences morality judgements (Corey et al., 2017; Costa, Foucart, Hayakawa, et al., 2014; Geipel, Hadjichristidis, & Surian, 2015a, 2015b, 2016), and reduces a well-described automatic self-bias (Ivaz, Costa, & Duñabeitia, 2015).

These foreign language influences have been discussed with respect to prominent dual-process accounts of decision making (Kahneman, 2003), which distinguish emotional and cognitive factors (Hayakawa, Costa, Foucart, &

[1]Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany
[2]Bernstein Center for Computational Neuroscience Berlin, Berlin, Germany
[3]Department of Neurological Surgery, University of California, San Francisco, California, CA, USA
[4]Center for Cognitive Neuroscience Berlin, Freie Universität Berlin, Berlin, Germany
[5]Institute for Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Corresponding author:
Christoph W Korn, University Medical Center Hamburg–Eppendorf, Martinistrasse 52, 20246 Hamburg, Germany.
Email: c.korn@uke.de

Keysar, 2016; Pavlenko, 2012). Foreign language use is thought to induce emotional distance and thereby alter the evaluation of affective information (Caldwell-Harris, 2014, 2015), although not all studies find straightforward support for this notion (Conrad, Recio, & Jacobs, 2011; Ponari et al., 2015). Differences in emotional processing seem more pronounced when the foreign language has been learned in a classroom setting rather than via immersion (Caldwell-Harris, 2015).

Motivated by these findings, we assessed here whether foreign language use reduces optimistic estimates about the personal future. Optimism is often regarded as one of the most important decision-making biases with a clear-cut emotional component (Sharot & Garrett, 2016; Shepperd et al., 2013; Taylor & Brown, 1988; Taylor, Collins, & Skokan, 1989). Research on optimism is often conducted from a clinical perspective because a lack of optimism, or even extreme pessimism, characterises depressive disorders (Disner, Beevers, Haigh, & Beck, 2011; Roiser, Elliott, & Sahakian, 2012).

All of the methods used to assess optimism rely—more or less explicitly—on the assumption that optimism only arises for events that bear an emotional connotation, such that decision makers overestimate the occurrence of positive events and underestimate the occurrence of negative events (Lefebvre, Lebreton, Meyniel, Bourgeois-Gironde, & Palminteri, 2016; Puri & Robinson, 2007; Scheier, Carver, & Bridges, 1994; Sharot & Garrett, 2016; Shepperd et al., 2013; Weinstein, 1980). Consequently, we assessed here whether participants differentially estimated their probabilities of experiencing affective life events (such as getting cancer or dying before reaching the age of 60) when using their native versus a foreign tongue. For Experiment 1, which was conducted in the laboratory, we adapted an updating task that has received considerable attention in the last years (Korn, Sharot, Walter, Heekeren, & Dolan, 2014; Sharot & Garrett, 2016; Sharot, Korn, & Dolan, 2011). In this task, participants first estimate their personal probabilities of experiencing negative events and then receive feedback about the baseline probabilities of these events in the population. This feedback can be desirable (negative event less probable in the population than estimated for the self) or undesirable (negative event more probable in the population than estimated for the self). On the basis of such feedback, participants can update their estimates about their personal future, and the degree of updating is measured by asking participants to re-estimate their personal probabilities a second time. Thus, we employ this task primarily to assess personal estimates per se and additionally to probe updating of these estimates. Compared with healthy controls, estimates are more pessimistic in patients suffering from depression (Korn et al., 2014) or borderline personality disorder (Korn, La Rosee, Heekeren, & Roepke, 2016). Regarding updating, it has been shown that healthy participants take desirable information about their personal future more into account than undesirable information (Kuzmanovic, Jefferson, & Vogeley, 2015; Ma et al., 2016; Sharot & Garrett, 2016; Sharot, Guitart-Masip, Korn, Chowdhury, & Dolan, 2012; Sharot et al., 2011), and this asymmetry is absent in depressive patients (Garrett et al., 2014; Korn et al., 2014) but not in borderline personality disorder patients (Korn et al., 2016).

In three online experiments (Experiments 2a, 2b, and 2c), we used a classic and well-replicated way of assessing optimism that relies on asking participants to estimate the probability of positive and/or negative life events happening to them and/or to an average person, who is similar to them (Shepperd et al., 2013). This is commonly referred to as the indirect method to assess comparative optimism at the group level, in the sense that participants do not directly rate how much more or less likely they are to experience an event in comparison with a relevant average person but instead participants give two separate ratings, one for themselves and one for an average person. For positive events, optimism is indicated by higher probability estimates for the self versus the other person. For negative events, the opposite is the case, which implies that optimism is indicated by a directional interaction for a larger difference between positive versus negative events for self- versus other-estimates. We expected that this interaction term, and specifically the self-estimates, would be reduced in a foreign language. Overall, this method has the great advantage that it allows easy and swift assessments with a few items, which makes the method ideal for online experiments and for using life events that can be understood by participants with low foreign language proficiency. We took advantage of online experiments to collect larger samples in Experiments 2a, 2b, and 2c, which allowed us to assess whether putative foreign language effects on optimism would be leveraged by foreign language proficiency. Online experiments have become a well-validated and well-received method of data collection that is integral to many branches of psychological research (Behrend, Sharek, Meade, & Wiebe, 2011; Chandler & Shapiro, 2016; Gosling & Mason, 2015; Rand, 2012; Stewart, Chandler, & Paolacci, 2017). We have taken care to follow the state-of-the-art in terms of quality standards.

Some studies have reported altered arousal ratings (or psychophysiological measurements) as indices of emotionality due to foreign language use (Pavlenko, 2012), but few studies have related emotion ratings to foreign language effects (Geipel et al., 2015a, 2015b, 2016; Hadjichristidis, Geipel, & Savadori, 2015). Here, we explored how arousal ratings relate to foreign language effects on optimism. Because language switching can affect cognitive control levels and thus influence decision biases (Oganian et al., 2016), no language switches occurred in the present experiments. Participants received the test material in either their native language German or in a foreign language (English or French). We made sure that participants were immersed in the respective language during the entire task.

Experiment 1: personal estimates and updating in the laboratory

Method

Participants. We recruited a total of 51 participants at Freie Universität Berlin, of which we included 45 for analyses (see below for exclusion criteria and Table 1 for demographics of the final sample). Participants gave informed consent and were paid. The study was approved by the Ethics Committee of the Charité—Universitätsmedizin Berlin. Participants performed the updating task in either their native language German (original $n=24$) or the foreign language English (original $n=27$) after an unrelated decision-making task that was conducted in the same language and is reported elsewhere (see pilot data in Korn, Heekeren, Oganian, 2018). Assignment of participants to the two groups was randomly determined by the time participants responded to our recruitment. After exclusion, the two groups were matched in age but differed in gender composition (Table 1). For completeness, we therefore performed and report additional analyses including the factor gender. To preview, gender had no significant influence.

To exclude switching effects, participants in the English group were informed beforehand that the experiment would take place in English only (under the supervision of a bilingual experimenter who addressed the participants exclusively in English). Participants completed a questionnaire on their language abilities, which included self-ratings of their English proficiency in reading, listening, writing, and speaking on a Likert-type scale from 1 (*single words*) to 7 (*native level*). To further assess proficiency, participants performed the Lexical Test for Advanced Learners of English (LexTALE, www.lextale.com; Lemhöfer & Broersma, 2012). As expected, mean self-rated proficiency correlated with the English LexTALE scores in the English group, Pearson's $r=0.459$, $p=.028$. In addition, participants completed a questionnaire assessing trait optimism (Life Orientation Test–Revised [LOT-R]; Scheier et al., 1994). The LOT-R includes items such as "Overall, I expect more good things to happen to me than bad" or "In uncertain times, I usually expect the best."

Updating task. In each trial of the first block of the task, participants were presented with negative events (e.g., cancer) for 2 s and were asked to estimate their probability of experiencing these within their lifetime (maximum of 6 s for response; see Supplementary Table 1 for the 70 events used here). After each personal estimate and a short fixation period (1 s), participants were presented with the baseline probability of the given event in the population (2 s). A new trial with a new event started after a short fixation (1 s). In the second block, participants re-estimated their personal probabilities, which allowed us to assess the updating of their estimates (timings as in the first block, except that no baseline probabilities were shown). To control for potential memory effects, after the second block, participants were asked to recall the baseline probabilities presented in the first block. In addition, participants rated all presented life events on seven subjective scales including arousal (see Table 1). The task was presented using the MATLAB (MathWorks, Natick, MA, USA) toolbox Cogent (http://www.vislab.ucl.ac.uk/cogent.php) on a standard PC.

We excluded events that the respective participants in the English group did not understand. To assess understanding, participants were presented with two lists at the end of the experiment: The first list comprised all life events in the English language version, and participants were asked to circle all words that they did not understand at all. On the second list, they saw all English items along with their German translation and were asked to circle events that they misunderstood. For this reason, slightly fewer items were included in the analyses of the English versus the German group (60.0 vs 64.1 items, see Table 1; in both groups, trials were excluded if participants did not answer in time or if participants' estimates were exactly the same as the population baseline, because in these cases feedback cannot be classified as desirable or undesirable; see next section).

Analyses. The analysis procedures followed previous reports (Korn et al., 2014; Sharot et al., 2011). For the main analyses, we compared mean overall first estimates between the two groups using a two-sample t test. The additional analyses regarding the updating of participants' estimates are somewhat more elaborate: Specifically, we separated events for which participants received desirable and undesirable information. Desirable information implies that participants' initial estimate in the first block was higher (i.e., more pessimistic) than the baseline probability, whereas undesirable information implies that the initial estimate was lower (i.e., more optimistic) than the baseline probability. To assess how much participants changed their estimates in response to the received information, we calculated the difference between first and second estimates. To normalise these changes with respect to the initial differences between participants' estimates and baseline probabilities, we calculated an update score by dividing the mean differences between first and second estimates by the mean differences between first estimates and baseline probabilities

$$\text{Update score} = \frac{\text{mean}\left(\text{first} - \text{second estimate}\right)}{\text{mean}\left(\text{first estimate} - \text{baseline probability}\right)}$$

Table 1. Optimistic belief updating in the laboratory Experiment 1: Characteristics of participants and task variables.

	L1 (German)	FL (English)	Effects comparing L1 and FL groups	*p* values comparing L1 and FL groups	Gender effects (*p* values for interaction effects of gender with L1 and FL groups)
Characteristics					
N (final sample)	22	23	—	—	—
Percentage female	0.81	0.47	**$\chi^2(1) = 5.67$**	**$p = .018$**	—
Age (years)	24.7 (5.0)	24.1 (3.5)	$t(43) = 0.47$	$p > .6$	—
LOT-R	16.1 (5.2)	16.3 (3.8)	$t(43) = -0.16$	$p > .8$	—
LexTALE	—	69.4 (10.4)	—	—	—
Mean self-rated proficiency in English	—	5.0 (0.8)	—	—	—
Task variables					
n trials included overall	64.1 (6.3)	60.0 (6.6)	**$t(43) = 2.36$**	**$p = .023$**	$p > .9$
1st estimates (in percentage rated probability)	36.6 (5.4)	35.5 (6.7)	$t(43) = 0.60$	$p > .5$	$p > .2$
Update—Desirable	0.56 (0.22)	0.50 (0.27)	$t(43) = 0.84$	$p > .4$	$p > .6$
Update—Undesirable	0.45 (0.25)	0.37 (0.24)	$t(43) = 1.00$	$p > .3$	$p > .4$
Memory error	12.8 (3.7)	13.9 (4.0	$t(43) = -0.92$	$p > .3$	$p > .4$
Arousal rating[a]	3.5 (0.8)	3.5 (0.8)	$t(43) = 0.32$	$p > .7$	$p > .1$
Negativity rating[a]	4.1 (0.8)	4.2 (0.8)	$t(43) = -0.26$	$p > .7$	$p > .6$
Experience rating[b]	1.4 (0.4)	1.6 (0.4)	$t(43) = -1.21$	$p > .2$	$p > .05$
Familiarity rating[a]	3.1 (0.7)	3.2 (0.7)	$t(43) = -0.18$	$p > .8$	$p > .9$
Vividness rating[a]	3.3 (0.6)	3.0 (0.6)	$t(43) = 1.52$	$p > .1$	$p > .1$
Controllability rating[a]	2.8 (0.7)	2.7 (0.6)	$t(43) = 0.32$	$p > .7$	$p > .06$

L1: native language; FL: foreign language English; LOTR: Life Orientation Test–Revised; LexTALE: Lexical Test for Advanced Learners of English; *SDs*: standard deviations.
Data are given as means (*SDs*). "Memory error" refers to the absolute difference between the presented baseline probability and the number of participants recalled. The mean probability of the life events occurring was 29.5% (17.0). The group difference in the proportion of female participants had no effect on our results (see also "Results" section of Experiment 1). Significant tests are marked in bold.
[a]From 1 = *not at all* to 6 = *very much*.
[b]From 1 = *never happened* to 6 = *happened very often*.

This update score should lie in the range of 0 (no update) to 1 (maximum update), otherwise it can be assumed that participants did not understand or follow the task because they either updated in the opposite direction than the information indicated (i.e., update score below 0) or "overshot" the information provided (i.e., update score above 1). We set two criteria for excluding participants: First, participants had to have understood at least 10 events per condition and 35 events overall. Second, update scores for both conditions had to lie between 0 and 1. Four participants in the English and two in the German group did not meet these criteria.

Results

Participants' probability estimates of experiencing negative events did not differ between the two language groups, $p > .5$. As gender composition differed between groups, $\chi^2(1) = 5.67$, $p = .018$, we tested for effects of gender. Gender had no influence on probability estimates, $p\text{s} > .2$. Our results replicated the typically observed pattern of updating but did not show an effect of language. Specifically, in a 2 (desirable vs undesirable information) × 2 (native vs foreign language) mixed analysis of variance (ANOVA) on update scores, the main effect of information was significant, $F(1, 43) = 5.53$, $p = .023$, $\eta_p^2 = .11$, indicating higher updating for desirable versus undesirable information (Table 1). The main effect of language and the interaction effect of information and language did not reach significance, both $p\text{s} > .2$. Adding gender as a factor in the ANOVA did not alter the results: The main effect of desirable versus undesirable information remained significant, $F(1, 41) = 5.87$, $p = .020$, $\eta_p^2 = .13$, but all other effects did not reach significance (all $p\text{s} > .1$).

In addition, none of the other measures differed significantly between the two language groups. This included memory scores, seven subjective event ratings, and trait optimism scores. Again, gender had no effect (see Table 1 for statistics).

Discussion

In sum, Experiment 1 revealed no effect of foreign language use on participants' estimates of experiencing a range of negative life events in the personal future. In addition, foreign language use did not alter updating of participants' estimates. What could be potential reasons for this lack of evidence (apart from a true lack of effect)? First, participants had a quite high proficiency (HP) in their foreign language English, which is typical for the population of German university students (see Table 1). Second, the employed sample size was based on previous studies using the updating task. Most previous studies on foreign language effects tested larger samples (but often using fewer items). Third, gender composition differed between the two groups and although we found no differences of gender on any measure, such differences might have limited the power to detect influences of foreign language use. To address these concerns, we conducted three online experiments with larger numbers of participants and more diverse proficiency levels.

Experiment 2a: comparative optimism online (within-participants design)

Method: Experiment 2a

Participants. Participants were recruited via the panel of a German online survey system (https://www.soscisurvey.de), which we had used previously (Korn, Ries, Schalk, Oganian, & Saalbach, 2017; Oganian et al., 2016). Participants in this panel participate out of interest and are not paid. Three panel administrators reviewed the questionnaire in detail regarding comprehensibility and feasibility and also gave suggestions for improving the questionnaire with respect to data interpretability and statistical analyses. Participants were recruited via email by the panel administrators, ensuring that panel members are not solicited too often. Studies using this panel are limited to a duration of 15 to 20 min (which is one reason why we refrained from running the setup of Experiment 1 using this panel).

For a summary of participants' characteristics, see Supplementary Table 2. We have previously reported data from the same sample on a different task, namely, framing scenarios (Oganian et al., 2016). Exclusion followed the same and additional criteria as in our previous report: (a) German only mother tongue, (b) current residence in a German-speaking country, (c) birth in a German-speaking country, (d) age between 18 and 60. In addition, participants were excluded if (e) their self-rated language proficiency in any of the two possible foreign languages did not exceed (*basic knowledge*, see next paragraph), (f) comments at the end of questionnaire indicated knowledge of the research questions, and (g) more than five items were missing for any of the two types of probability estimates (see next section). Due to these strict criteria, out of the initial 927 participants who completed the questionnaire, 706 were included in the current analyses.

To recruit a large pool of participants with a large variability in foreign language proficiency, the foreign language condition was administered in either English or French (as these are the most often learned languages in school in German-speaking countries). In the beginning of the questionnaire, participants were asked whether they had ever learnt English and/or French and consequently estimated their proficiency in reading, listening, writing, and speaking on Likert-type scales from 1 (*single words*) to 7 (*native level*). If their mean self-rated proficiency was above 2 (*basic knowledge*), they were randomly assigned to the respective foreign language version or to the German version (i.e., participants whose mean proficiency in both English and French did not exceed 2 were not allowed to continue with the questionnaire or were excluded from all analyses; Supplementary Table 2). Thus, participants' proficiency in the foreign language tested ranged from above 2 to 7. Average foreign language proficiency (across English and/or French) did not differ between participants in the foreign and native language conditions, $ps < .9$.

Participants' level of education was assessed on a Likert-type scale according to the education system in German-speaking countries from 1 (*dropped out of school*) to 8 (*university degree*). Proficiency in the tested language did not significantly correlate with level of education or age, both $ps > .2$. Proficiency differed slightly but significantly between the two genders, $t(701) = 2.42$, $p = .016$ (female: $M = 5.0$, standard deviation [SD] = 1.8; male: $M = 5.3$, $SD = 1.8$; three persons did not indicate their gender). We therefore explicitly tested for effects of gender. To preview our results, we did not find any significant influence of gender.

Comparative optimism task: within-participants design. Participants estimated the probability of five positive and five negative life events by typing the percentage into an empty box (for stimuli, see Supplementary Table 3). They estimated the probability for themselves and for an average, other person on different screens. The exact wording presented to participants for the other person was: "An average other person participating in this study (same gender and age as you)." The order of the self/other conditions and the order of the items on the screen were randomised. Participants then rated the life events with respect to arousal (from 1 = *not arousing at all* to 6 = *very arousing*) and valence (from 1 = *very negative* to 6 = *very positive*). Participants had the option to indicate that they do not know the word (thus as in Experiment 1, we relied on participants to indicate whether they understood the words).

Analysis. We analysed participants' probability estimates and their arousal and valence ratings using linear mixed effects models (LMEs) as implemented in the R package lme4 (Baayen, Davidson, & Bates, 2008; http://cran.r-project.org/web/packages/lme4/index.html). We used LMEs because they allow including random effects for

Table 2. Experiment 2: Comparative optimism online—Mean ratings.

Experiment	Proficiency group	*N*		Ratings of event probability in percentage (*SD*)[a]					Arousal ratings[b,c]		Valence ratings[b,d]	
		Self-condition	Other-condition	Self-positive	Self-negative	Other-positive	Other-negative	Comparative optimism (interaction)[e]	Positive	Negative	Positive	Negative
2a (self–other condition within participants)	L1	241	[f]	43.9 (15.7)	32.3 (13.3)	35.1 (10.8)	35.6 (12.0)	12.0	52.4 (18.4)	58.8 (19.2)	76.5 (16.1)	22.3 (12.9)
	HP	154	[f]	42.1 (15.1)	31.7 (11.8)	35.8 (12.2)	36.1 (12.4)	10.8	50.7 (18.4)	50.6 (21.1)	75.4 (13.8)	24.3 (12.2)
	LP	311	[f]	40.5 (15.5)	35.2 (15.4)	35.6 (13.8)	36.2 (14.1)	5.8	46.9 (21.0)	52.0 (22.3)	70.2 (17.7)	23.4 (13.0)
2b (self–other condition between participants)	L1	127	128	59.4 (16.6)	46.1 (16.4)	46.2 (13.2)	48.9 (12.8)	16.0	46.2 (19.4)	58.6 (16.3)	78.8 (10.7)	20.7 (9.8)
	HP	50	57	60.6 (15.0)	46.6 (14.1)	46.4 (14.0)	44.3 (15.2)	12.0	51.5 (18.1)	52.1 (19.0)	81.3 (8.6)	22.1 (9.7)
	LP	87	81	53.0 (18.8)	49.1 (17.8)	47.5 (16.5)	50.1 (17.0)	6.5	46.5 (19.3)	48.8 (19.9)	73.6 (14.3)	24.4 (11.9)
2c (self–other condition within participants)	L1	224	[f]	56.9 (17.9)	46.7 (17.2)	45.6 (11.8)	50.2 (15.2)	14.9	44.3 (19.3)	63.7 (17.5)	75.6 (15.4)	23.0 (13.0)
	HP	91	[f]	55.4 (15.4)	44.9 (15.4)	43.6 (12.2)	44.9 (13.5)	11.9	50.7 (17.8)	51.8 (23.2)	78.1 (12.2)	22.9 (9.4)
	LP	158	[f]	53.4 (19.6)	47.2 (17.6)	47.2 (16.0)	49.1 (17.3)	8.1	49.6 (21.6)	55.5 (21.9)	71.3 (19.9)	25.2 (15.0)

L1: native language (i.e., mean proficiency of 7 on a 7-point scale); HP: high proficiency (proficiency equal or above 5); LP: low proficiency (proficiency below 5 and equal or above 2); *SD*: standard deviation.
Data are given as counts, percentages, or means (*SD*).
[a]In Experiment 2a, participants were asked to type in their ratings of the lifetime probability of the event occurring. In Experiments 2b and 2c, they were asked to indicate their ratings on a Visual Analogue Scale with a slider. This difference in response format likely explains the overall differences in ratings between the samples. [b]In Experiments 2a and 2c, all participants gave arousal *and* valence ratings. In Experiment 2b, participants gave either arousal or valence ratings (independent of the assignment to the self- or other-condition). [c]Arousal ratings were given on a scale from 1 = *low* to 100 = *high*. [d]Valence ratings were given on a scale from 1 = *negative* to 100 = *positive*. [e]Comparative optimism is given by the following interaction term: (self-positive–self-negative) – (other-positive–other-negative).
[f]Self/other was manipulated within subject.

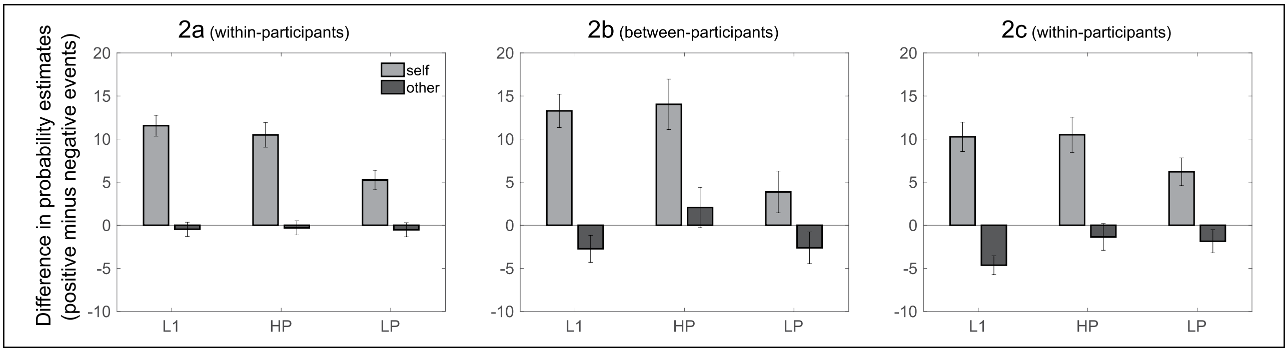

Figure 1. Illustration of reduced comparative optimism for low foreign language proficiency. In three online experiments (Experiments 2a, 2b, and 2c), comparative optimism was lower in the low-proficiency (LP) group than in the high proficiency group (HP) and in the native language group (L1). This was indicated by the significant interaction of proficiency with the optimism term (self-positive–self-negative) – (other-positive–other-negative). The bar graphs illustrate two patterns: First, self-estimates for positive events were generally higher than self-estimates for negative events. In contrast, other-estimates were relatively similar for positive and negative events. Second, LP, in particular, changed self-ratings. Error bars depict standard errors of the mean (based on participant-wise mean differences between probability estimates for positive vs negative events).

both participants and items. In line with widely accepted recommendations (Barr, 2013), we included random slopes for the highest order interaction term of within-unit factors. We tested for effects of foreign language proficiency by including a continuous variable that consisted of the mean self-rated proficiency of the language in which participants completed the questionnaire (with mother tongue coded as 7). The LME for the probability ratings is given by the following formula according to the conventions of lme4 in R: ratings = self–other × positive–negative × proficiency + (1 + self–other × positive–negative | participant) + (1 + self–other + proficiency | item). In addition, proficiency was coded as a categorical variable with three levels to allow comparison with our previous report (Oganian et al., 2016): L1 (*native language*), HP (mean proficiency of 5 and higher), and low proficiency (LP; mean proficiency below 5; see Supplementary Table 4). To facilitate comparison of our results with the existing literature on comparative optimism (Shepperd et al., 2013), we conducted a mixed ANOVA on the means of participants' probability estimates per condition: 2 (self–other) × 2 (positive–negative) × 3 (proficiency groups: L1-HP-LP). In contrast to the corresponding LME, this ANOVA neglects within-participants variance and the random effect of items.

We analysed the relationship between probability estimates and arousal ratings (as well as valence ratings). To test this relationship, we calculated straightforwardly interpretable correlations across all items for each participant (i.e., these analyses assess the relationship between arousal ratings and probability estimates at the level of each individual participant and do therefore not directly correspond to a mediation analysis across participants). To compare the strengths of these correlations across participants, participant-wise Pearson's r values were transformed to Fisher's z values to obtain normally distributed values (which was confirmed by visual inspection of the resulting distribution).

Results: Experiment 2a

Proficiency and probability estimates. Comparative optimism is operationalised as an interaction in how participants estimate the probability of experiencing positive and negative events (factor: positive–negative) for themselves versus an average person (factor: self–other). Comparative optimism implies that participants estimate their personal probability for positive events higher than for negative events—but show a reduced or absent difference between positive versus negative events in probability estimates for the average person. Consequently, we performed a linear mixed effects analysis, in which comparative optimism is characterised by positive values for the interaction term (self-positive–self-negative) – (other-positive–other-negative). Proficiency was coded as a continuous variable. This linear mixed effects analysis revealed a significant triple interaction of self versus other estimates × positive versus negative events × proficiency, $p < 10^{-4}$, indicating reduced comparative optimism at lower proficiency levels (see Tables 2 and 3). The conceptually same triple interaction emerged when treating proficiency as a categorical variable, $p < 10^{-3}$ (Table 2 and Supplementary Table 4), and when conducting a mixed ANOVA, $F(2, 703) = 10.02$, $p < 10^{-4}$, $\eta^2_p = .028$. Figure 1 provides a graphical illustration of these effects. The intra-individual *SD*s across all probability estimates per participant did not significantly correlate with the proficiency of the used language, $p > .3$, which makes it unlikely that participants used the scale in starkly different ways.

As proficiency differed between the two genders, $t(701) = 2.42$, $p = .016$, we ran an additional linear mixed effects analysis, which included the factor gender.

Table 3. Experiment 2: Comparative optimism online—Linear mixed effects models on ratings of event probability (with proficiency as a continuous variable).

Effect	2a (self–other condition within participants)[a]		2b (self–other condition between participants)[a]		2c (self–other condition within participants)[a]	
	Estimates	*t*-values	Estimates	*t*-values	Estimates	*t*-values
Intercept	**35.9**	**5.45**	**53.0**	**8.82**	**47.0**	**6.63**
Self–other	−0.8	−0.43	−4.4	−0.85	−0.7	−0.21
Positive–negative	−3.4	−0.36	−4.9	−0.59	2.2	0.22
Proficiency	−0.6	−1.27	−1.1	−1.43	−0.1	−0.14
Self–other × positive–negative	−0.9	−0.33	−2.0	0.31	−1.4	−0.29
Self–other × proficiency	**0.7**	**2.35**	1.0	1.25	0.5	1.32
Positive–negative × proficiency	**2.2**	**3.53**	**2.7**	**2.67**	1.2	1.32
Self–other × positive–negative × proficiency	**−1.6**	**−4.04**	**−2.5**	**−2.41**	**−2.0**	**−3.17**

Significant *t*-values are marked in bold. See Supplementary Table 4 for conceptually analogous analyses with proficiency as a categorical variable, which also indicated a reduction of comparative optimism with lower proficiency.
[a]In all three experiments, the triple interaction self–other × positive–negative × proficiency level was significant, indicating that the magnitude of comparative optimism (self–other × positive–negative) was reduced with lower proficiency.

Gender had no significant influence and the triple interaction of self versus other estimates × positive versus negative events × proficiency remained significant (Supplementary Table 5).

The HP group included a higher percentage of participants tested in English versus French (i.e., 0.76 of participants tested in English, Supplementary Table 2), while the LP group was relatively more balanced (i.e., 0.35 of participants tested in English). We therefore ran an additional linear mixed effects analysis that only compared the L1 and LP groups but not the HP group. The triple interaction of self versus other estimates × positive versus negative events × proficiency held, $p < .0005$. In contrast, this triple interaction was not significant when only comparing the L1 and HP groups, $p > .5$.

In addition, we tested the influences both of proficiency (as continuous covariate) and of the three used languages within the same mixed ANOVA on mean probability estimates per condition (i.e., this ANOVA included data from all participants). The triple interaction including proficiency was significant, $F(1,\ 702) = 8.96$, $p = .003$, $= .13$, but the triple interaction with the factor language was not significant, $p > .8$. Thus, foreign language proficiency influenced comparative optimism, and it is unlikely that this depended on specific effects of English or French as foreign languages.

Proficiency and arousal ratings. Motivated by the research of foreign language effects on emotional processing (Caldwell-Harris, 2014, 2015; Pavlenko, 2012), we specifically assessed the effects of proficiency on arousal ratings. Indeed, lower proficiency was associated with lower arousal ratings (significant main effect of proficiency; see Table 4 for full regression model). For completeness, we also explored the effects of proficiency on valence ratings and found that lower proficiency was associated with smaller difference in valence ratings between positive and negative events (interaction of positive vs negative events × proficiency, see Table 4). Similarly as for the probability estimates, the intra-individual *SD*s across arousal or valence ratings per participant did not significantly correlate with the proficiency of the used language, both $ps > .2$. These results suggest an influence of language proficiency on emotionality ratings.

Relationship between probability estimates and arousal ratings. Based on the previous findings, we asked the follow-up question: Are the observed foreign language effects on comparative optimism related to the reduced emotional arousal elicited by the stimuli? To test this, we correlated each participant's probability estimates with their arousal ratings across the 10 items separately for self/other estimates. (For this analysis, arousal ratings for negative items were scored negatively such that more extreme arousal ratings for negative events would relate to lower probability estimates.) Across all participants, the relationship between estimates and ratings was stronger for self- than for other-estimates—mean arousal ratings for self $r = .24$, t test against zero: $t(667) = 14.60$, $p < 10^{-9}$, Cohen's $d_z = 0.57$; mean arousal ratings for other $r = .03$, t test against zero: $t(669) = 1.78$, $p = .076$; one-sample t test comparing these correlations for self and other across participants: $t(656) = 13.56$, $p < 10^{-9}$, $d_z = 0.53$ (Table 5). (A similar pattern emerged in exploratory analyses for valence ratings: self $r = .39$, other $r = .20$, see Table 5.) Thus, arousal ratings were associated with self-estimates, such that higher arousal related to more extreme probability estimates for self-estimates but not for other-estimates.

Language proficiency correlated with the strength of the relationship between self-estimates and arousal ratings

Table 4. Experiment 2: Comparative optimism online—Linear mixed effects models on arousal and valence ratings.

Rating type	Effect	2a (self–other condition within participants)[a]		2b (self–other condition between participants)[b]		2c (self–other condition within participants)[a]	
		Beta-estimates	*t*-values	Beta-estimates	*t*-values	Beta-estimates	*t*-values
Arousal ratings	Intercept	**48.6**	**9.38**	**42.4**	**6.80**	**47.2**	**7.48**
	Positive–negative	−1.7	−0.24	5.4	0.63	4.8	0.55
	Proficiency	**1.2**	**2.58**	**2.1**	**3.11**	**2.1**	**3.75**
	Positive–negative × proficiency	−0.5	−0.92	**−2.2**	**−2.44**	**−2.9**	**−4.33**
Valence ratings	Intercept	**25.6**	**5.70**	**28.3**	**5.11**	**27.7**	**5.49**
	Positive–negative	**42.3**	**6.52**	**38.5**	**4.75**	**39.9**	**5.36**
	Proficiency	−0.4	−1.53	**−1.1**	**−2.78**	**−0.8**	**−2.28**
	Positive–negative × proficiency	**1.7**	**3.64**	**3.0**	**4.54**	**2.1**	**3.54**

All models are given by the following formula according to the conventions of R: ratings = positive–negative × proficiency + (1 + positive–negative | participant) + (1 | item). Proficiency was inserted as a continuous variable. Significant *t*-values are marked in bold.
[a]In line with our hypothesis based on previous demonstrations of reduced arousal in foreign language settings, the main effects of proficiency were significant for arousal ratings in both experiments. In Experiments 2a and 2c, all participants rated both arousal and valence. [b]In Experiment 2b, participants gave either arousal or valence ratings (independent of the assignment to the self- or other-condition).

Table 5. Experiment 2: Comparative optimism online—Relation between probability estimates, emotionality ratings, and language proficiency.

Rating type	Effect	2a (self–other condition within participants)[a]			2b (self–other condition between participants)[b]			2c (self–other condition within participants)[a]		
		Participant-wise Pearson's *r* between estimates and ratings[c]	Correlation of participant-wise *r* and language proficiency		Participant-wise Pearson's *r* between estimates and ratings[c]	Correlation of participant-wise *r* and language proficiency		Participant-wise Pearson's *r* between estimates and ratings[c]	Correlation of participant-wise *r* and language proficiency	
			r	*p*		*r*	*p*		*r*	*p*
Arousal ratings	Self	.24	**.12**	**<.005**	.27	**.22**	**<.05**	.20	.02	>.5
	Other	.03	−.03	>.4	.00	.01	>.8	**−.03**	**−.14**	**<.01**
Valence ratings	Self	.39	.00	>.8	.40	.14	>.1	.37	.04	>.4
	Other	.20	**−.16**	**<.001**	.25	−.01	>.8	.14	**−.13**	**<.01**

Bold font indicates tests that survived Bonferroni correction for two tests. Similar results were observed when using Spearman's ρ instead of Pearson's *r*: In Experiments 2a and 2b, participant-wise ρ values for self-estimates and arousal ratings correlated with proficiency; Experiment 2a, $\rho = .12$, $p < .01$; Experiment 2b, $\rho = .25$, $p < .005$. In Experiment 2c, this relationship did not reach significance, $p > .7$.
[a]In Experiments 2a and 2c, all participants made estimates for self and other and rated both arousal and valence. [b]In Experiment 2b, participants either made estimates for self or other and gave either arousal or valence ratings (independent of the assignment to the self- or other-condition).
[c]Participant-wise *r* values were transformed to Fisher's *z* values before averaging and back-transformed for the mean values provided.

such that participants with higher proficiency had a tighter relationship between rated arousal and estimated probability for self, $r = .12$, $p < .005$ (Bonferroni correction for two tests; one for self and one for other). This effect indicates that using a low-proficient foreign language entails a looser coupling between estimating the probability of experiencing events and the arousal elicited by contemplating these events.

In summary, this high-powered online experiment with a within-participants design showed that using a foreign language with LP reduces comparative optimism. Further analyses suggested that LP weakens the relationship between self-estimates and arousal.

Experiment 2b: comparative optimism online (between-participants design)

To provide sound evidence for an influence of low foreign language proficiency on comparative optimism, we aimed at replicating the effects of Experiment 2a in a between-participants design.

Method: Experiment 2b

We recruited an independent sample (via the same panel and online survey system), which performed the comparative optimism task after unrelated tasks (to be reported elsewhere). Using the same exclusion criteria as for Experiment 2a, we retained 530 of the 616 participants who completed the survey (see Supplementary Table 2). The procedure was the same as in Experiment 2a with two exceptions: First and most importantly, Experiment 2b employed a between-participants design: Participants estimated probabilities either for themselves or for an average, other person. Independently, they rated either arousal or valence. This reduced the testing time for each participant. Second, participants estimated the probability of the life events using a slider with the anchors 0% and 100% (the exact chosen number was indicated numerically above the slider). Average foreign language proficiency (across English and/or French) did not differ between participants in the foreign and native language conditions, *p*s < .9. Mean proficiency in the tested language did not significantly correlate with level of education or age, nor did it differ according to gender, all *p*s > .2.

Analyses were similar to Experiment 2a. The random-effects structure of the LMEs was adapted to the between-participants design: ratings = self-other × positive–negative × proficiency + (1 + positive–negative | participant) + (1 + self–other + proficiency | item).

Results: Experiment 2b

Proficiency and probability estimates. Experiment 2b replicated the effects of Experiment 2a. Comparative optimism was reduced with lower proficiency (Figure 1 and Table 2) as indicated by a significant triple interaction, $p = .017$ (Table 3). The same was true when proficiency was treated as a categorical variable, $p = .048$ (Table 2 and Supplementary Table 4), and a trend for the triple interaction emerged in the less powerful mixed ANOVA, $F(2, 524) = 2.94$, $p = .054$, $\eta_p^2 = .011$.

As in Experiment 2a, the HP group included a higher percentage of participants tested in English versus French (i.e., 0.84, Supplementary Table 2). In contrast, the LP group had almost equal proportions of participants tested in the two foreign languages (i.e., 0.49 of participants tested in English). When comparing only the L1 and LP groups, the triple interaction was still significant, $p = .014$. But as in Experiment 2a, the triple interaction was not significant when comparing the L1 and HP groups, $p > .2$. Again, intra-individual *SD*s across all probability estimates per participant did not significantly correlate with the proficiency of the used language, $p = .075$.

Proficiency and arousal ratings. Lower proficiency was again related to lower arousal ratings (and also to a reduced valence difference between positive and negative events; Table 4). For arousal ratings, intra-individual *SD*s per participant did not significantly correlate with the proficiency of the used language, $p > .8$. For valence ratings, the corresponding correlation was significant but rather small, $r = .15$, $p = .016$.

Relationship between probability estimates and arousal ratings. As in Experiment 2a, we tested the relationship between language proficiency, estimates of event probability, and arousal ratings (Table 5). Again the relationship between estimates and ratings was stronger for self- than for other-estimates in the between-participants design—mean arousal ratings for self $r = .27$, *t* test against zero: $t(138) = 7.99$, $p < 10^{-9}$, $d_z = 0.67$; mean arousal ratings for other $r = .00$, *t* test against zero: $p > .9$; two-sample *t* test comparing these correlations for self and other across participants: $t(265) = 5.72$, $p < 10^{-7}$, Cohen's $d_s = 0.70$ (Table 5).

The same effect as in Experiment 2a emerged with respect to proficiency: Across participants, the item-wise correlations between self-estimates and arousal ratings were reduced with lower proficiency, $r = .22$, $p < .05$ (Bonferroni correction for two tests: one for self and one for other).

Taken together, Experiment 2b replicated the effects of Experiment 2a in a between-participants design.

Experiment 2c: comparative optimism online (within-participants design)

Following the suggestions of anonymous reviewers, we conducted an additional replication experiment to exclude that the observed effects of low foreign language proficiency on comparative optimism were driven by other potentially relevant factors, such as individual differences in trait optimism, socio-economic status, immigrant status, parents' educational level, trait and state anxiety, need for cognition, motivation, and healthy lifestyle. The selection of these variables was to a large degree guided by research testing for cognitive effects of bilingualism (Antón et al., 2014; Duñabeitia et al., 2014; Gathercole et al., 2014; Paap & Sawi, 2014; Paap, Sawi, Dalibar, Darrow, & Johnson, 2014).

In an effort to enhance transparency and quality, we preregistered all relevant details with the Open Science Framework (https://osf.io/98j6m/) and performed an a priori power analysis on the basis of Experiment 2a.

Method: Experiment 2c

Participants. Participants were recruited using the same panel and online survey system as in Experiments 2a and 2b. Applying the same exclusion criteria as for Experiments 2a and 2b, we retained 473 of the 545 participants who completed the relevant sections of the survey.

Design. Experiment 2c was intended as a direct replication of Experiment 2a with the additional assessment of a variety of control measures (see below). Therefore, Experiment 2c followed the same within-participants design as

Experiment 2a (with the only technical difference that estimates and ratings were given on a slider as in Experiment 2b and not typed in as in Experiment 2a). Participants completed the comparative optimism task followed by the assessment of various metrics (see below) and unrelated tasks (to be reported elsewhere).

As before, mean proficiency in the tested language did not differ according to gender, $p\text{s} > .3$, nor did it significantly correlate with the level of participants' own education, $p = .08$. Surprisingly, mean proficiency in the tested language correlated significantly with age, $r = -.16$, $p < .001$, which was not the case in the two previous online experiments. To preview, control analyses showed that age was unrelated to comparative optimism. Again, average foreign language proficiency (across English and/or French) did not differ between participants in the foreign and native language conditions, $p\text{s} < .1$.

Power analysis. Sample size considerations were based on an a priori power analysis using the software program G*Power. Our goal was to obtain .85 power at the standard .05 alpha error probability using a non-directional hypothesis for the triple interaction in the mixed ANOVA on the means of participants' probability estimates per condition: 2 (self–other) × 2 (positive–negative) × 3 (proficiency groups: L1-HP-LP). On the basis of the effect size observed in Experiment 2a, we obtained a targeted sample size of at least 432 participants.

Control measures. To control for potential individual differences that might relate to comparative optimism and foreign language proficiency, we collected the following measures in addition to participants' gender, age, and education: (a) trait optimism as assessed by the LOT-R (Scheier et al., 1994), (b) socio-economic status as assessed by monthly income (from 1 = *less than €250* to 9 = *more than €5,000*), (c) the educational level of the parent with the higher educational level, (d) participants' state and trait anxiety as assessed by the 10-item short versions of the State-Trait Anxiety Inventory (Laux, Glanzmann, Schaffner, & Spielberger, 1981), (e) participants' need for cognition as assessed using the four-item version (Bless, Wänke, Bohner, Fellhauer, & Schwarz, 1994), (f) participants' motivation to fill out the questionnaire (*How motivated were you to answer the preceding questions?* From 1 = *not at all* to 5 = *very much*), (g) their current stress level (*How often do you feel stressed?* From 1 = *never* to 5 = *very often*), and their propensities to (h) eat healthy food (*How often do you eat healthy food?* From 1 = *never* to 5 = *very often*) and to (i) do sports (*How often do you do sports?* From 1 = *never* to 5 = *very often*).

As in Experiments 2a and 2b, participants were excluded if they spoke any other mother tongue except for German and if they were born outside a German-speaking country. This rules out that participants themselves were immigrants. In Experiment 2c, we additionally asked whether any parent was born outside a German-speaking country, which was the case for 5.9% of participants.

Results: Experiment 2c

Proficiency and probability estimates. The respective results of the preregistered Experiment 2c replicated those of Experiments 2a and 2b. Lower proficiency was related to reduced comparative optimism (Table 2), as demonstrated by significant triple interactions in LMEs treating proficiency as continuous variable, $p = .002$ (Table 3), or as categorical variable, $p = .034$ (Supplementary Table 4). In mixed ANOVAs, the corresponding triple interactions were significant as well—proficiency as continuous variable: $F(1, 471) = 14.19$, $p < .001$, $\eta_p^2 = .029$; proficiency as categorical variable: $F(2, 470) = 4.93$, $p = .008$, $\eta_p^2 = .021$. The latter mixed ANOVA constitutes the primary preregistered analysis procedure. All following analyses should be regarded as supplementary with respect to the preregistration.

As participants' age correlated significantly with their mean proficiency in the used language, we tested an LME with the factor age instead of proficiency, but no significant effects of age emerged, that is, all absolute *t*-values were below 1.2 and specifically the triple interaction of self versus other estimates × positive versus negative events × age was not significant, $p\text{s} > .2$. In addition, we conducted a mixed ANOVA with proficiency as categorical variable to which we added age as a covariate. The effects of age were not significant, all $p\text{s} > .1$, and the triple interaction of self versus other estimates × positive versus negative events × proficiency was significant, $F(2, 469) = 5.50$, $p = .004$, $\eta_p^2 = .023$. That is, proficiency but not age was related to comparative optimism.

Again, the HP group but not the LP group included a higher percentage of participants tested in English versus French (i.e., HP: 0.81; LP: 0.47, Supplementary Table 2). A significant triple interaction of self versus other estimates × positive versus negative events × proficiency emerged, when comparing only the L1 and LP groups, $p = .010$, but not when comparing the L1 and HP groups, $p > .3$.

In addition, a mixed ANOVA on all participants with proficiency (as continuous covariate) and a factor coding for the three languages revealed a significant triple interaction with the covariate proficiency, $F(1, 469) = 5.62$, $p = .018$, $\eta_p^2 = .12$, but not with the factor language, $p > .8$. There was no significant correlation between the intra-individual *SD*s across all probability estimates per participant and the proficiency of the used language, $p > .2$.

Control measures. Across participants, none of the nine control measures listed above correlated significantly with the independent variable, that is, mean proficiency in the used language, all ps > .3 (even without performing Bonferroni correction). That is, participants can be considered as matched with regard to these nine measures. Also, mean proficiency in the used language did not differ between participants whose parents were both born in a German-speaking country and those participants with at least one parent born outside a German-speaking country, $p > .5$.

Proficiency and arousal ratings. The results reported in Table 4 show that lower proficiency was associated with lower arousal ratings (and also with a reduced valence difference between positive and negative events). Intra-individual *SD*s across arousal or valence ratings per participant did not significantly correlate with the proficiency of the used language, both ps > .1.

Relationship between probability estimates and arousal ratings. The relationship between estimates and ratings was more pronounced for self- than for other-estimates—mean arousal ratings for self $r = .20$, t test against zero: $t(467) = 11.04$, $p < 10^{-9}$, $d_z = 0.51$; mean arousal ratings for other $r = -.03$, t test against zero: $p = .074$; one-sample t test comparing these correlations for self and other across participants: $t(466) = 11.95$, $p < 10^{-9}$, $d_z = 0.55$ (Table 5). In Experiment 2c, the item-wise correlations between self-estimates and arousal ratings were not significantly related to proficiency, $r = .02$, $p > .5$, which might indicate that the sample size of Experiment 2c was insufficient to detect this relationship.

Discussion: Experiments 2a, 2b, and 2c

All three online experiments indicated lower comparative optimism for participants with lower proficiency in a foreign language. These online experiments allowed us to obtain data from participants across a wide range of foreign language proficiencies on the employed 7-point scale (25, 50, and 75 percentiles: Experiment 2a: 3.00, 4.00, and 5.25; Experiment 2b: 3.25, 4.50, and 5.25; Experiment 2c: 3.00, 4.00, and 5.25 for participants who performed the task in a foreign language). This is a larger spread than in our initial laboratory experiment (25, 50, and 75 percentiles: 4.06, 5.25, and 5.50) and might be one possible reason why we observed significant influences of foreign language use in the online samples but not the laboratory sample. The notion that LP users are specifically affected by foreign language use is bolstered by our findings that the LP but not the HP groups differed from the L1 groups in Experiments 2a, 2b, and 2c.

Participants were randomly assigned to the native and foreign language conditions (given that they had sufficient proficiency in the respective foreign language). In all experiments and especially in Experiment 2c, we tested whether individual differences in the language used, that is, the independent variable, were matched with respect to a variety of control variables. This was generally the case. The only exceptions were differences in gender in Experiment 2a and differences in age in Experiment 2c. Importantly, however, the fact that these differences in gender or age were not related to comparative optimism and the fact that each of these differences only occurred in one of three experiments rule out that they affected our general findings.

General discussion

The aim of the current report was to test the role of foreign language use as a factor that alters estimates about the personal future, which often appear optimistically biased. Three independent online experiments revealed that LP in a foreign language reduced comparative optimism at the group level. That is, low foreign language proficiency reduced the commonly observed tendency to estimate one's personal future as brighter than that of an average person (Shepperd et al., 2013). Our findings link this effect of foreign language proficiency to emotional arousal: Overall, arousal ratings correlated with estimates for the personal future and these correlations were smaller for lower proficiency levels in two of the three experiments. This tentatively suggests that personal estimates are coupled to the subjective emotional arousal elicited by thinking about possible future life events.

The recent debate that study replications do not always produce the same results as the initial studies (Open Science Collaboration, 2015) and reviewer comments motivated us to provide upfront replications of our results: one replication with a modification in methodology (i.e., within-participants design in the original experiment vs between-participants design in the replication) and one preregistered replication with the assessment of additional control metrics.

In contrast to the significant effects of foreign language use on comparative optimism (in three online experiments), we found no significant evidence for effects of foreign language use on self-estimates for negative events or on updates of these self-estimates (in a laboratory experiment). We hold it likely that this was—at least in part—due to differences in foreign language proficiency. The three online experiments suggest that the foreign language effect is only present in LP but not in HP individuals. Participants in the laboratory experiment had rather high foreign language proficiency. In addition, the laboratory task employed a different methodology: Only self-estimates for experiencing negative life events were elicited, and not self- and other-estimates for positive and negative events as in the online experiments. Therefore, it is possible that questions about negative and/or positive events contributed to the different results across experiments. We deem it therefore an interesting avenue for future research to investigate specifically whether LP foreign language use affects optimism differentially for positive and negative events. Furthermore, participants in the laboratory experiment received information about the population baseline probability of these events on a trial-by-trial basis so that we could additionally assess how they updated their estimates in response to this information. The amount of overall updating did not differ between the foreign and the native language groups, and we found stronger updating for desirable than undesirable information in both groups.

Although a recent report has questioned whether optimistic updating extends to positive events and whether it results from neglecting the base rates of the events (Shah, Harris, Bird, Catmur, & Hahn, 2016), others have provided evidence that biased updating is robust and that the raised concerns are not supported by additional data (Garrett & Sharot, 2014, 2017; Kuzmanovic, Jefferson, & Vogeley, 2016). Also, please note that a group comparison of participants' first estimates is in any case unrelated to the raised concerns, which solely targeted the analyses of the asymmetric updating pattern. Relatedly, we encourage further investigations of less commonly used assessment methodologies, for example, reward learning tasks or estimating life expectancy (Lefebvre et al., 2016; Puri & Robinson, 2007; Shepperd et al., 2013).

Our current findings are in line with theories on bilingualism that stress emotional processing as the mechanistic link between foreign language use and its effects on decision making and judgement (Caldwell-Harris, 2014, 2015; Pavlenko, 2012). Our observation that low foreign language proficiency leads to diminished comparative optimism dovetails with the emotional distance account (Hayakawa et al., 2016; Pavlenko, 2012), an interpretation favoured in several previous reports on diverse tasks such as risky decision making (Costa, Foucart, Arnon, et al., 2014; Keysar et al., 2012), speeded self-evaluations (Ivaz et al., 2015), and morality judgements (Costa, Foucart, Hayakawa, et al., 2014; Geipel et al., 2015a, 2015b, 2016). We demonstrate an impact of low foreign language proficiency on subjective arousal, which is in line with the common notion of proficiency as one of the potential candidates for conferring emotionality to a foreign tongue (Caldwell-Harris, 2014, 2015; Pavlenko, 2012). An upshot is that highly proficient speakers of a foreign language might attain native-like levels of emotional processing and thus show similar behaviour regardless of the language they are currently using (as observed in all experiments reported here).

A rather large and diverse literature has discussed why holding optimistic beliefs may be adaptive: Optimism may lead to adaptive strategies such as exploring unknown options, coping with adverse events, or engaging in precautionary behaviour (Alicke & Sedikides, 2009; Haselton et al., 2009; Haselton & Nettle, 2006; Johnson & Fowler, 2011; Nettle & Bateson, 2012; Puri & Robinson, 2007; Solberg Nes & Segerstrom, 2006). It is thus an interesting avenue for future research to investigate whether foreign language use affects such behavioural strategies.

At a superficial glance, our findings seem to diverge from those of a recent study that assessed judgements of risk and benefit for various hazards (e.g., food preservatives, nuclear power plants, and climate change) in native Italian speakers with English as a foreign language (Hadjichristidis et al., 2015). In that study, judgements were overall more positive under foreign language use than in the native tongue. However, participants in the study by Hadjichristidis et al. were asked to judge the impact of hazards to the society at large—and not to estimate the probability of such events occurring to themselves or an average person. We deem it an interesting question for future research to directly compare the effects of foreign language use on probability estimates as well as risk and benefit judgements for personal and public events within the same participants.

Although emotional distance may be the most prominently debated explanation for foreign language effects (Hayakawa et al., 2016; Pavlenko, 2012), other cognitive influences are likely to play a role. Importantly, it seems possible that at LP levels, the experimental materials may not be processed to the depth that they can be at higher or native proficiency levels. Effects of processing depth could, for example, be tested by manipulating the allotted time or the presentation of the experimental materials—in conjunction to foreign language use (cf. Korn et al., 2017, for effects related to presenting framing scenarios in hard-to-read fonts). We have recently shown that the cognitive costs entailed by switching between languages reduce the framing effect for hypothetical scenarios (Oganian et al., 2016). Notably, our previous study did not find evidence for an influence of language proficiency on the framing effect. It is thus possible that proficiency alters emotional but not cognitive factors involved in foreign language effects. Optimism about the self and framing effects can both be regarded as decision biases with emotional components, but otherwise they differ considerably. We hold it likely that foreign language use affects decision biases via different routes depending on the specific type of bias.

A simple dichotomy between emotional and cognitive factors seems useful for summarising initial findings but falls short of explaining subtle contextual effects. Fine-grained delineations of foreign language effects on different types of decision biases will hopefully contribute to understanding why decision biases exist at all. This implies in turn that foreign language effects offer a great tool for dissecting the causes of decision biases—independent of whether the biases are adaptive or not in the given situation (Hayakawa et al., 2016). Specifically, theories on optimism seem to assume that the emotional processing associated with imagining one's personal future drives behaviour towards adaptive strategies (Alicke & Sedikides, 2009; Haselton et al., 2009; Haselton & Nettle, 2006; Puri & Robinson, 2007; Solberg Nes & Segerstrom, 2006).

In conclusion, our findings establish that estimates about the personal future become less optimistic when made in a low-proficient foreign language. This highlights the role of foreign language use—especially at LP—for optimism and decision biases in general.

Acknowledgements

We thank Lioba La Rosée and Ulrike Schlickeiser for help with data acquisition. C.W.K., Y.O., and H.R.H. planned the experiments and wrote the manuscript. C.W.K. and Y.O. collected and analysed the data.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by a grant from the Deutsche Forschungsgemeinschaft (GRK 1598/1). Y.O. was supported by a research training grant from the German Research Council (GRK 1891/1). C.W.K. was supported by the SFB TRR 169 from the German Research Council.

Supplementary material

The Supplementary Material is available at: qjep.sagepub.com.

ORCID iD

Yulia Oganian https://orcid.org/0000-0002-3451-8938

References

Alicke, M. D., & Sedikides, C. (2009). Self-enhancement and self-protection: What they are and what they do. *European Review of Social Psychology*, *20*(1), 1–48. doi:10.1080/10463280802613866

Antón, E., Duñabeitia, J. A., Estévez, A., Hernández, J. A., Castillo, A., Fuentes, L. J., & Carreiras, M. (2014). Is there a bilingual advantage in the ANT task? Evidence from children. *Frontiers in Psychology*, *5*, 398. doi:10.3389/fpsyg.2014.00398

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412. doi:10.1016/j.jml.2007.12.005

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology*, *4*, 328. doi:10.3389/fpsyg.2013.00328

Behrend, T. S., Sharek, D. J., Meade, A. W., & Wiebe, E. N. (2011). The viability of crowdsourcing for survey research. *Behavior Research Methods*, *43*, 800–813. doi:10.3758/s13428-011-0081-0

Bless, H., Wänke, M., Bohner, G., Fellhauer, R. F., & Schwarz, N. (1994). Need for Cognition: Eine Skala zur Erfassung von Engagement und Freude bei Denkaufgaben [Need for cognition: a scale measuring engagement and happiness in cognitive tasks]. *Zeitschrift Für Sozialpsychologie*, 147–154. doi:10.1111/j.1467-839X.2008.00269.x

Caldwell-Harris, C. L. (2014). Emotionality differences between a native and foreign language: Theoretical implications. *Frontiers in Psychology*, *5*, 1–4. doi:10.3389/fpsyg.2014.01055

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. doi:10.1177/0963721414566268

Chandler, J., & Shapiro, D. (2016). Conducting clinical research using crowdsourced convenience samples. *Annual Review of Clinical Psychology*, *12*(1), 53–81. doi:10.1146/annurev-clinpsy-021815-093623

Conrad, M., Recio, G., & Jacobs, A. M. (2011). The time course of emotion effects in first and second language processing: A cross cultural ERP study with German-Spanish bilinguals. *Frontiers in Psychology*, *2*, 1–16. doi:10.3389/fpsyg.2011.00351

Corey, J. D., Hayakawa, S., Foucart, A., Aparici, M., Botella, J., & Costa, A. (2017). Our moral choices are foreign to us. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*, 1109–1128.

Costa, A., Foucart, A., Arnon, I., Aparici, M., & Apesteguia, J. (2014). "Piensa" twice: On the foreign language effect in decision making. *Cognition*, *130*(2), 236–254. doi:10.1016/j.cognition.2013.11.010

Costa, A., Foucart, A., Hayakawa, S., Aparici, M., Apesteguia, J., Heafner, J., & Keysar, B. (2014). Your morals depend on language. *PLoS ONE*, *9*(4), e94842. doi:10.1371/journal.pone.0094842

Disner, S. G., Beevers, C. G., Haigh, E. A. P., & Beck, A. T. (2011). Neural mechanisms of the cognitive model of depression. *Nature Reviews Neuroscience*, *12*, 467–477. doi:10.1038/nrn3027

Duñabeitia, J., Hernández, J., Antón, E., Macizo, P., Estévez, A., Fuentes, L., & Carreiras, M. (2014). The inhibitory advantage in bilingual children revisited: myth or reality? *Experimental Psychology*, *61*, 234–251. doi:10.1027/1618-3169/a000243

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*, 5983–5989. doi:10.1523/JNEUROSCI.3622-14.2015

Garrett, N., & Sharot, T. (2014). How robust is the optimistic update bias for estimating self-risk and population base rates? *PLoS ONE*, *9*(6), e98848. doi:10.1371/journal.pone.0098848

Garrett, N., & Sharot, T. (2017). Optimistic update bias holds firm: Three tests of robustness following Shah et al. *Consciousness and Cognition*, *50*, 12–22. doi:10.1016/j.concog.2016.10.013

Garrett, N., Sharot, T., Faulkner, P., Korn, C. W., Roiser, J. P., & Dolan, R. J. (2014). Losing the rose tinted glasses: Neural substrates of unbiased belief updating in depression. *Frontiers in Human Neuroscience*, *8*, 639. doi:10.3389/fnhum.2014.00639

Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Guasch, N. V., & Jones, L. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks. *Frontiers in Psychology*, *5*, 1–14. doi:10.3389/fpsyg.2014.00011

Geipel, J., Hadjichristidis, C., & Surian, L. (2015a). How foreign language shapes moral judgment. *Journal of Experimental Social Psychology*, *59*, 8–17. doi:10.1016/j.jesp.2015.02.001

Geipel, J., Hadjichristidis, C., & Surian, L. (2015b). The foreign language effect on moral judgment: The role of emotions and norms. *PLoS ONE*, *10*(7), e0131529. doi:10.1371/journal.pone.0131529

Geipel, J., Hadjichristidis, C., & Surian, L. (2016). Foreign language affects the contribution of intentions and outcomes to moral judgment. *Cognition*, *154*, 34–39. doi:10.1016/j.cognition.2016.05.010

Gosling, S. D., & Mason, W. (2015). Internet research in psychology. *Annual Review of Psychology*, *66*, 877–902. doi:10.1146/annurev-psych-010814-015321

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*, 117–129. doi:10.1037/xap0000044

Haselton, M. G., Bryant, G. A., Wilke, A., Frederick, D. A., Galperin, A., Frankenhuis, W. E., & Moore, T. (2009). Adaptive rationality: An evolutionary perspective on cognitive bias. *Social Cognition*, *27*, 733–763. doi:10.1521/soco.2009.27.5.733

Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review*, *10*(1), 47–66. doi:10.1207/s15327957pspr1001

Hayakawa, S., Costa, A., Foucart, A., & Keysar, B. (2016). Using a foreign language changes our choices. *Trends in Cognitive Sciences*, *20*, 791–793. doi:10.1016/j.tics.2016.08.004

Ivaz, L., Costa, A., & Duñabeitia, J. A. (2015). The emotional impact of being myself: Emotions and foreign-language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 489–496. doi:10.1037/xlm0000179

Johnson, D., & Fowler, J. (2011). The evolution of overconfidence. *Nature*, *477*, 317–320.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *The American Psychologist*, *58*, 697–720. doi:10.1037/0003-066X.58.9.697

Keysar, B., Hayakawa, S. L., & An, S. G. (2012). The foreign-language effect: Thinking in a foreign tongue reduces decision biases. *Psychological Science*, *23*, 661–668. doi:10.1177/0956797611432178

Korn, C.W., Heekeren, H.R., & Oganian, Y. (2018). The framing effect in a monetary gambling task is robust in minimally verbal language switching contexts. *Quarterly Journal of Experimental Psychology*. Advance online publication. doi: 10.1177/1747021818769259.

Korn, C. W., La Rosee, L., Heekeren, H. R., & Roepke, S. (2016). Processing of information about future life events in borderline personality disorder. *Psychiatry Research*, *246*, 719–724. doi:10.1016/j.psychres.2016.07.067

Korn, C. W., Ries, J., Schalk, L., Oganian, Y., & Saalbach, H. (2017). A hard-to-read font reduces the framing effect in a large sample. *Psychonomic Bulletin and Review*, *25*, 696–703. doi:10.3758/s13423-017-1395-4

Korn, C. W., Sharot, T., Walter, H., Heekeren, H. R., & Dolan, R. J. (2014). Depression is related to an absence of optimistically biased belief updating about future life events. *Psychological Medicine*, *44*, 579–592. doi:10.1017/S0033291713001074

Kühberger, A. (1998). The influence of framing on risky decisions: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *75*(1), 23–55. doi:10.1006/obhd.1998.2781

Kuzmanovic, B., Jefferson, A., & Vogeley, K. (2015). Self-specific optimism bias in belief updating is associated with high trait o optimism. *Journal of Behavioral Decision Making*, *28*, 281–293. doi:10.1002/bdm.1849

Kuzmanovic, B., Jefferson, A., & Vogeley, K. (2016). The role of the neural reward circuitry in self-referential optimistic belief updates. *NeuroImage*, *133*, 151–162.

Laux, L., Glanzmann, P., Schaffner, P., & Spielberger, C. D. (1981). *STAI—State-Trait-Angstinventar. Theoretische Grundlagen und Handanweisung* [STAI-State-Trait-Anxiety-Inventory. Theoretical bases and instructions.] Weinheim, Germany: Beltz Test GmbH. (in German)

Lefebvre, G., Lebreton, M., Meyniel, F., Bourgeois-Gironde, S., & Palminteri, S. (2016). Asymmetric reinforcement learning: Computational and neural bases of positive life orientation. *Biorxiv*. doi:10.1101/038778

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*, 325–343. doi:10.3758/s13428-011-0146-0

Ma, Y., Li, S., Wang, C., Liu, Y., Li, W., Yan, X., & Han, S. (2016). Distinct oxytocin effects on belief updating in response to desirable and undesirable feedback. *Proceedings of the National Academy of Sciences*, *113*, 9256–9261. doi:10.1073/pnas.1604285113

Nettle, D., & Bateson, M. (2012). The evolutionary origins of mood and its disorders. *Current Biology*, *22*, R712–R721. doi:10.1016/j.cub.2012.06.020

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*, 140–148. doi:10.1037/xlm0000161

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. doi:10.1126/science.aac4716

Paap, K. R., & Sawi, O. (2014). Bilingual advantages in executive functioning: Problems in convergent validity, discriminant validity, and the identification of the theoretical constructs. *Frontiers in Psychology*, *5*, 962. doi:10.3389/fpsyg.2014.00962

Paap, K. R., Sawi, O. M., Dalibar, C., Darrow, J., & Johnson, H. A. (2014). The brain mechanisms underlying the cognitive benefits of bilingualism may be extraordinarily difficult to discover. *AIMS Neuroscience*, *1*, 245–256. doi:10.3934/Neuroscience.2014.3.245

Pavlenko, A. (2012). Affective processing in bilingual speakers: Disembodied cognition? *International Journal of Psychology*, *47*, 405–428. doi:10.1080/00207594.2012.743665

Ponari, M., Rodríguez-Cuadrado, S., Vinson, D., Fox, N., Costa, A., & Vigliocco, G. (2015). Processing advantage for emotional words in bilingual speakers. *Emotion*, *15*, 644–652.

Puri, M., & Robinson, D. T. (2007). Optimism and economic choice. *Journal of Financial Economics*, *86*(1), 71–99. doi:10.1016/j.jfineco.2006.09.003

Rand, D. G. (2012). The promise of Mechanical Turk: How online labor markets can help theorists run behavioral experiments. *Journal of Theoretical Biology*, *299*, 172–179.

Roiser, J. P., Elliott, R., & Sahakian, B. J. (2012). Cognitive mechanisms of treatment in depression. *Neuropsychopharmacology*, *37*, 117–136. doi:10.1038/npp.2011.183

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A re-evaluation of the Life Orientation Test. *Journal of Personality and Social Psychology*, *67*, 1063–1078.

Shah, P., Harris, A. J. L., Bird, G., Catmur, C., & Hahn, U. (2016). A pessimistic view of optimistic belief updating. *Cognitive Psychology*, *90*, 71–127. doi:10.1016/j.cogpsych.2016.05.004

Sharot, T. (2011). *The optimism bias*. New York, NY: Pantheon Books.

Sharot, T., & Garrett, N. (2016). Forming beliefs: Why valence matters. *Trends in Cognitive Sciences*, *20*(1), 25–33. doi:10.1016/j.tics.2015.11.002

Sharot, T., Guitart-Masip, M., Korn, C. W., Chowdhury, R., & Dolan, R. J. (2012). How dopamine enhances an optimism bias in humans. *Current Biology*, *22*, 1477–1481. doi:10.1016/j.cub.2012.05.053

Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience*, *14*, 1475–1479.

Shepperd, J. A., Klein, W. M. P., Waters, E. A., & Weinstein, N. D. (2013). Taking stock of unrealistic optimism. *Perspectives on Psychological Science*, *8*, 395–411.

Solberg Nes, L., & Segerstrom, S. C. (2006). Dispositional optimism and coping: A meta-analytic review. *Personal and Social Psychology Review*, *10*, 235–251.

Stewart, N., Chandler, J., & Paolacci, G. (2017). Crowdsourcing samples in cognitive science. *Trends in Cognitive Sciences*, *21*, 736–748. doi:10.1016/j.tics.2017.06.007

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, 193–210. doi:10.1037//0033-2909.103.2.193

Taylor, S. E., Collins, R. L., Skokan, L. A., & Aspinwall, L. G. (1989). Maintaining positive illusions in the face of negative information: Getting the facts without letting them get you. *Journal of Social and Clinical Psychology*, *8*, 114–129.

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

Weinstein, N. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, *39*, 806-820.

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*, 427-436. doi:10.1080/20445911.2016.1139583