

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23

Does bilingualism shape inhibitory control in the elderly?

Eneko Antón, Yuriem Fernández García, Manuel Carreiras and Jon Andoni Duñabeitia

Basque Center on Cognition, Brain and Language (BCBL); Donostia, Spain

Contact information:

Eneko Antón

Paseo Mikeletegi 69, 2

20009 Donostia – Spain

0034943309300

e.anton@bcbl.eu

Abstract

Bilingualism has been argued to benefit executive functioning. However, recent research suggests that this advantage may stem from uncontrolled factors or incorrectly matched samples. In this study we test the effects of bilingualism on elderly lifelong bilinguals whose cognitive abilities are in decline, thus making any benefits more salient. Firstly we compare 24 bilinguals and 24 carefully matched monolinguals on verbal and the numerical Stroop tasks, obtaining no differences in monitoring or inhibitory measures. Secondly we explore the modulations that the proficiency in the L2 might cause to executive control functions, as measured by the same tasks, by testing 70 elderly bilinguals who vary in their L2 mastery from very low to perfectly fluent. Results show no modulation in any of the indices due to L2 proficiency. These results add to the growing body of evidence showing that the bilingual advantages might indeed be due to other factors rather than bilingualism.

Keywords: Bilingualism; bilingual advantage; lifelong bilingualism; executive functioning.

Does bilingualism shape inhibitory control in the elderly?

The impact of bilingualism in human cognition has become a popular topic in the past several years and the debate about the so-called bilingual advantage is driving the attention of many researchers in the field. The hypothesis of this apparent advantage of bilinguals over monolinguals in some domain-general non-linguistic skills is grounded on the assumption that bilingualism provides speakers with an additional cognitive training: when it comes to facing communicative situations with different speakers and different demands, bilinguals have to choose between the two languages that they have available in their repertoire to make their communication efficient and to adequate it to the requirements of the current context. Given that the languages that a bilingual speaker speaks are active always (e.g., Lagrou, Hartsuiker, & Duyck, 2013; Midgley, Holcomb, van Heuven, & Grainger, 2008; Thierry, & Wu, 2007), the language that is not meant to be used in a given situation has to be inhibited to avoid any cross-linguistic contamination that could harm effective communication. In contrast, monolinguals do not face these situations, so no language choice is needed and no language inhibition is required. Based on this difference between bilinguals and monolinguals, the hypothesis of the bilingual advantage postulates that bilinguals have increased experience with their executive functions, which is a cognitive construct formed by inhibition (i.e., the ability to suppress dominant responses), shifting (i.e., the ability to switch between tasks), and monitoring (i.e., the ability to update the information in the working memory; see Miyake et al., 2000; Miyake & Friedman, 2012). In other words, given that bilinguals are constantly monitoring their linguistic environment and inhibiting the non-relevant language, the so-called hypothesis of the bilingual advantage suggests that they have enhanced executive functions as compared to monolinguals (e.g., Bialystok, Craik, Green, & Gollan, 2009). In fact, evidence in favor of this advantage has been provided in a variety of studies using non-linguistic executive control tasks, mainly (but not only) focusing on two relevant measurements: inhibitory and monitoring skills (e.g., Bialystok, Craik, Klein & Viswanathan, 2004; Bialystok, Craik & Luk, 2008, 2012; Hernández, Costa, Fuentes, Vivas, & Sebastián-Gallés, 2010; Kroll & Bialystok, 2013).

On the one hand, some researchers have argued that, considering that bilinguals need to reliably inhibit the non-target language, this capacity to avoid intrusion of inappropriate pieces of information would constitute the enhanced skill. Consequently, the advantage should be easily captured by tasks measuring inhibitory skills, such as the classic flanker task (Eriksen & Eriksen, 1974), the Stroop task (Stroop, 1935) or the Simon task (Simon & Rudell, 1967). All these tasks share some features: there are incongruent and congruent trials. Incongruent trials present irrelevant and distracting information and support a different or opposing response from the one expected.

1 This irrelevant information must either be ignored or the unwanted response needs to be inhibited.
2 Congruent trials are those in which the irrelevant information favors the same response as the one
3 that the relevant information does. Crucially, the difference between the average responses to these
4 two kinds of conditions (i.e., the conflict effect) is found to be smaller in bilinguals (i.e. better
5 inhibitory skills) in some of these tasks (Bialystok, Craik, Klein, & Viswanathan, 2004; Costa,
6 Hernández, & Sebastián-Gallés, 2008; Bialystok, Craik, & Luk, 2008). These differences have been
7 interpreted as a better ability of bilinguals to deal with irrelevant and confusing information and to
8 suppress it, as they do in their daily life when they suppress then non-target language.

9 On the other hand, the purported bilingual advantage has been also grounded on bilinguals'
10 enhanced monitoring skills (Costa et al. 2009), which stems from the constant need of bilinguals to
11 keep track of the communicative context and monitoring the language needed for the
12 communicative situation in question. These demands would enhance the bilinguals' skill to
13 efficiently tackle taxing contexts in which different stimuli (some relevant, some irrelevant) are
14 present and intermixed, similarly to the everyday situations in which they have to pay attention to
15 different stimuli (the linguistic representations), and select only the relevant ones. In the last few
16 years some researchers have closely related monitoring abilities with global reaction times (RTs) in
17 tasks involving different kinds of congruency conditions, and thus a bilingual advantage in
18 monitoring has been defended empirically by research showing that bilinguals are overall faster
19 than monolinguals when the context of the task is high demanding and that, crucially, this
20 advantage vanishes if the context is low demanding (Bialystok, 2006; Costa et al., 2009; Costa &
21 Sebastián-Gallés, 2014; but see Paap, Johnson, & Sawi, 2014 for a discussion on the impurity of the
22 use of global RTs as a measure of monitoring).

23 In contrast to the accounts that favor a genuine enhancement of cognitive control as a
24 consequence of bilingualism, a strong and increasing line of research suggests that the so-called
25 bilingual advantage might be due to hidden demographic factors that tend to be differently
26 distributed among the bilingual and monolingual populations under study (e.g., immigrant status,
27 educational level or socio-economic status, among many others; see Paap, Johnson, & Sawi,
28 2015a), and that these factors, and not bilingualism *per se*, may be responsible for the observed
29 differences. This counter-argument, which constitutes one of the building blocks of opponents to
30 the so-called bilingual advantage, started with the pioneering work by Morton and Harper (2007).
31 Being aware of the crucial role played by many demographic factors in the development and
32 mastery of executive functioning (among many other cognitive skills), they pointed out that
33 preceding research comparing groups with different linguistic contexts had completely neglected
34 the role of these factors. Some years later, other researchers also brought this same idea to the

1 debate, and the importance of possibly confounding socio-demographic factors started to be
2 acknowledged (see Hilchey & Klein, 2011). Good examples of this are provided by socioeconomic
3 status (SES), which has been often correlated with a better performance in executive functioning
4 tasks, as well as by the immigrant status and other ethnicity-related factors, which are known to
5 affect the quality and speed of the performance on executive function tasks (e.g., Mezzacappa,
6 2004; Noble, Norman, & Farah, 2005). Bearing this in mind, one can easily observe how some of
7 these factors have been overlooked in many studies reporting a bilingual advantage. For instance,
8 Bialystok and Martin (2004) compared Canadian monolingual and bilingual children without
9 measuring SES. Bialystok and Shapero (2005) compared Canadian monolinguals with immigrant
10 bilinguals coming from different linguistic and ethnic background without measuring SES. In this
11 same line, Bialystok, Craik and Luk (2008) mainly included immigrants in their sample of
12 bilinguals (20 out of 24 individuals). More recently, Engel de Abreu and colleagues (Engel de
13 Abreu, Cruz-Santos, Tourinho, Martin, and Bialystok, 2012) matched their monolinguals and
14 bilinguals for SES, but they took each sample from a different country (Portugal and Luxemburg,
15 respectively). All these studies yielded a bilingual advantage, which the authors unambiguously
16 attributed to bilingualism instead of considering the potential effects of the abovementioned factors.
17 Along these lines, in some extensive reviews, Paap and colleagues (see Paap & Greenberg, 2013;
18 Paap, Johnson, & Sawi, 2015a, 2015b) identified several other deficiencies that may have
19 contribute to the appearance of a bilingual advantage in preceding studies, including not only Type I
20 errors due to inadequately matched groups or small sample sizes, but also uncontrolled external
21 factors or task-dependent effects.

22 Crucially, when the mentioned confounding factors are controlled for and participants are
23 carefully matched, any sign for a bilingual advantage, either in inhibition (reduced conflict effect)
24 or in monitoring (overall reduced RTs) tends to vanish (see Paap & Greenberg, 2013). In this line,
25 Morton and Harper (2007) tried to replicate the findings obtained by Bialystok, Craik, Klein and
26 Viswanathan (2004) using the Simon task in children, but as opposed to Bialystok et al., they
27 matched both groups on SES, immigrant status and ethnicity. They found no bilingual advantage in
28 this carefully-controlled experimental setting. Following this same rationale, null results (no
29 bilingual advantage) have been replicated in the last several years in studies in which participant
30 groups were carefully matched for the factors mentioned above and in which the number of
31 participants tested was relatively high (see Antón et al., 2014; Duñabeitia et al., 2014; Paap &
32 Greenberg, 2013; Paap, Johnson, & Sawi, 2015a, 2015b; Gathercole et al., 2014).

33 While methodological concerns might capture the heterogeneity of the results presented so
34 far in studies comparing monolingual and bilingual young adults, an additional critical issue related

1 to the development of certain cognitive skills of interest has been recently raised. Research in this
2 field has mainly focused on studying populations formed by young adults, which are in normal
3 terms at the peak of their domain-general cognitive abilities (around 20-40 years of age, see
4 Hartshorne & Germine, 2015). This fact makes them likely to show a “ceiling effect” moment with
5 respect to their domain-general cognitive abilities, which can mask or hide the presence of a
6 potential difference between bilinguals and monolinguals, and thus the bilingualism-driven
7 differences might arguably be hard to capture.

8 Therefore, if any relevant cognitive differences are to arise as a consequence of bilingualism,
9 they would be expected to be most salient when the cognitive skills are not at their maximum, but
10 rather when they are still developing (childhood) or already declining (elderly). This is precisely
11 what has been found by Bialystok and her colleagues (Bialystok, Martin, & Viswanathan, 2005):
12 they administered the Simon task to 5 year-old children, young adults (20 years of age), middle-
13 aged adults (30-59 years of age) and older adults (60-80 years), and found that the bilingual
14 advantage was present in children and middle-aged and older adult groups, but absent in the young
15 adult participants. However, the general picture does not seem to be consistent in these samples
16 either, and the evidence for a bilingual advantage in the childhood and in the elderly has been also
17 questioned recently. While some studies show an advantage for bilingual children as compared to
18 their monolingual peers (e.g., Martin-Rhee & Bialystok, 2008; Yang, Yang, & Lust, 2011), recent
19 findings suggest that, using a carefully matched and large sample, the advantage disappears in
20 children samples, as it is the case with young adults too (Antón et al. 2014; Duñabeitia et al., 2014;
21 Gathercole et al. 2014).

22 The scenario with bilingual seniors, whose cognitive abilities are already declining due to
23 normal cognitive deterioration caused by age, is also unclear and foggy. There are a relatively
24 reduced number of studies providing seemingly consistent evidence favoring a behavioral
25 advantage in tasks measuring different forms of cognitive control and executive functions in elderly
26 bilinguals. For example, Bialystok, Craik and Luk (2008) showed stronger differences between
27 bilinguals and monolinguals in older samples as compared to young adults in both the Simon and
28 the verbal Stroop tasks (see also Bialystok, Craik, Klein, & Viswanathan, 2004; Gold, Kim,
29 Johnson, Kryscio, & Smith, 2013), but lately other researchers have suggested that those pieces of
30 evidence are not entirely reliable and replicable. For instance, Kirk and colleagues (Kirk, Fiala,
31 Scott-Brown, & Kempe, 2014) tried to replicate the findings obtained by Bialystok et al. (2004)
32 demonstrating a reduction of the Simon effect in older bilinguals, and found neither signs of a
33 bilingual advantage in inhibition nor a difference in global reaction times (i.e., no advantage in
34 monitoring). This inconsistent pattern is also observable in other studies testing elderly participants,

1 where no group differences appear (see Kousaie & Phillips, 2012; see also de Bruin, Bak & Della
2 Sala, 2015, for a study comparing monolinguals and active and passive bilinguals in the Simon
3 Arrow task with no evidence for any bilingual advantage).

4 Furthermore, defenders of the bilingual advantage have also argued in favor of a different
5 kind of advantage in the elderly by considering bilingualism a form of protection against dementia
6 (i.e., the neuro-protective value of bilingualism; see Bialystok, Abutalebi, Bak, Burke, & Kroll,
7 2016, for a recent review). For example, Bialystok and colleagues (Bialystok, Craik, & Freedman,
8 2007) analyzed records of patients that underwent a process of being diagnosed and treated for
9 dementia, and found that bilinguals manifested the incipient symptoms around four years later than
10 monolinguals, with no change in the later rate of progression or course of the illness. However,
11 these same results have not been fully replicated in other samples, and an increasing number of
12 authors deny that the symptomatology of neurodegenerative diseases is delayed in bilinguals due to
13 their seemingly greater cognitive reserve (e.g., Costa & Sebastian-Gallés, 2014; Paap, Johnson &
14 Sawi, 2015a; 2015b). As pointed out by Chertkow and colleagues (Chertkow et al., 2010), 90% of
15 the bilingual patients investigated by Bialystok et al. (2007) were immigrants, and in a replication
16 of these results they found critical interactions between immigrant status and bilingualism. In an
17 attempt to further clarify the scenario, Gollan and her colleagues (Gollan, Salmon, Montoya, &
18 Galasko, 2011) defended that the neuro-protective role of bilingualism is indeed real, but only in
19 people with low educational level. They found that higher degrees of bilingualism were associated
20 with a delay in the diagnosis, but only in bilinguals with a low educational level. As a potential
21 solution for the conundrum presented by the unclear debate of the role of bilingualism in patients
22 with dementia, a handful of studies have opted for a longitudinal approach, testing cohorts starting
23 from a baseline stage in which no signs of dementia are evident. Most of the studies using this
24 approach have shown no consistent delay in the onset of the symptoms caused by bilingualism
25 (Crane et al., 2009; Lawton, Gasquoine, & Weimer, 2015; Sanders, Hall, Katz, & Lipton, 2012;
26 Yeung, St. John, Menec, & Tyas, 2014; Zahodne, Schofield, Farrell, Stern, & Manly, 2014).

27 This contradictory evidence draws an unclear picture of the potentially beneficial effects that
28 lifelong bilingualism may have in elderly bilinguals in both executive functions and
29 neurodegenerative processes. According to the general view, it seems reasonable to think that if the
30 use of two languages yields any boosting or enhancement of domain-general cognitive abilities, this
31 should be easily captured in elderly lifelong bilinguals, since they have been actively exposed to
32 two languages for decades, consequently training their inhibitory and monitoring skills for much
33 longer than bilingual children or young adults. However, the evidence in this regard is far from
34 clear and additional studies are required to better understand whether or not bilingualism yields

1 beneficial effects for cognitive reserve in the elderly. The current study was designed to shed light
2 on this precise issue.

3 It has been suggested that instead of focusing on comparisons between bilinguals and
4 monolinguals (i.e., taking bilingualism as an “all or nothing” factor), bilingualism should be better
5 treated as a continuous rather than dichotomous factor. For example, Singh and Mishra (2013)
6 tested high and low proficient Hindi-English bilinguals on a modified saccadic arrow Stroop task
7 and found a group effect, indicating that high proficient bilinguals were faster and had reduced
8 conflict indices, thus showing that L2 proficiency can apparently modulate monitoring and conflict
9 resolution skills. However, some other researchers had also considered this hypothesis and found no
10 differences in inhibitory control, monitoring or switching based on participants’ L2 proficiency
11 (see, for example, Paap, Johnson, & Sawi, 2014). If the effect of the proficiency of the L2 found by
12 Singh and Mishra (2013) is robust and consistent, one should be able to easily replicate a similar
13 finding in groups of bilinguals immersed in a society in which the two languages they speak are
14 present and official by exploring how varying degrees of L2 proficiency modulate differently
15 classic markers of inhibitory skills. This was precisely the aim of our second set of experiments.

16 Here we present evidence from verbal & non-verbal Stroop tasks aimed at exploring if
17 elderly bilingual speakers differ from their monolingual peers in their inhibitory skills and/or in
18 their monitoring abilities; and if among bilinguals there is a modulation based on their L2
19 proficiency. Considering that many studies have consistently reported that bilinguals suffer a
20 disadvantage in production of spoken responses (Ivanova & Costa, 2007; Gollan, Montoya,
21 Fennema-Notestine, & Morris, 2005) we collected data from both the verbal Stroop task, i.e., the
22 classic color-naming version of the Stroop paradigm (Stroop, 1935) as well as from a less
23 linguistically-charged version of the same paradigm, the numerical Stroop task (Besner & Coltheart,
24 1979; Kaufmann et al., 2005; Santens & Verguts, 2011). These two paradigms were implemented
25 along the lines that Duñabeitia et al. (2014) followed in their study testing the reliability of the
26 bilingual advantage during childhood, where they found reliable differences between conditions
27 (i.e., the Stroop, congruency and incongruity indices or effects), which were *not* modulated by the
28 language background of the participants. In the current study, in Experiment 1A, we compared
29 monolingual and carefully matched non-immigrant highly proficient lifelong bilingual seniors in the
30 verbal Stroop task, while in the Experiment 1B we compared the same set of participants in the
31 numerical Stroop task. In addition to this, and considering that it has been suggested that
32 bilingualism should be treated as a continuous rather than dichotomous factor (e.g., Singh &
33 Mishra, 2013), in Experiment 2 we explored the impact of the “degree of bilingualism”, here
34 measured as non-immigrant bilinguals’ general proficiency in their second language, in executive

1 control functions. A wide range of bilingual seniors with varying degrees of second language (L2)
2 proficiency were tested in the verbal Stroop (Experiment 2A) and the numerical Stroop tasks
3 (Experiment 2B), to see if the classic markers of inhibition (as well as of monitoring) were
4 modulated by participants' L2 knowledge and mastery.

5

6 **Experiments 1A and 1B: Effect of lifelong bilingualism in executive control.**

7 In this first set of experiments, the hypothesis of the bilingual advantage in inhibitory skills
8 was tested by comparing elderly bilinguals' and monolinguals' performance in a verbal
9 (Experiment 1A) and in a numerical version (Experiment 1B) of the Stroop paradigm. Both groups
10 of participants were carefully matched in the potentially confounding factors discussed above (cf.
11 Paap & Greenberg, 2013), including general intelligence, socioeconomic status, immigration and
12 ethnicity, so that the only relevant difference between both groups was their linguistic profile.
13 Additionally, bearing in mind that some researchers have circumscribed the presence of a bilingual
14 advantage to lowly educated samples (see Gollan et al., 2011), most of the participants that were
15 recruited had a relatively low educational level (see below for details). If the previously reported
16 instances of bilingual advantages in similar tasks were not the result of the confounding factors that
17 we controlled for, both groups of participants would show significant differences either in the
18 magnitude of the Stroop effects (i.e., bilinguals should show a reduced Stroop effects as a reflection
19 of their enhanced inhibitory skills) or in the overall reaction times (with bilinguals performing
20 overall faster due to their better monitoring abilities).

21

Method

22 **Participants**

23 48 seniors (28 females) were recruited in the Basque Country (mean age 69.06, SD=4.56; age
24 range = 61-78). All of them reported to have normal or corrected-to-normal vision and none of
25 them had any history of chronic neuropsychological disorders. Every participant signed a written
26 consent form approved by the Ethics and Research Committees of the Basque Center on Cognition,
27 Brain and Language (BCBL) establish.

28 Half of the participants (14 female; mean age= 68.75, SD=4.42) were Spanish monolinguals;
29 and the other half (14 female, mean age of 69.38, SD=4.58) were Basque-Spanish bilinguals who
30 use both languages everyday and rate themselves as highly proficient in comprehension and
31 production in Basque (average score of 8.04 (SD=0.95) where 1 is really poor level and 10 is
32 perfectly fluent) as well as in Spanish (average score of 8.67 (SD=1.17) over 10). Apart from self-

1 rated proficiency, which has been extensively used in the literature and reported to accurately
2 account for the actual proficiency (see, among many others, Clark, 1981; Heilenman, 1990;
3 LeBlanc and Painchaud, 1985), a native speaker of both languages interviewed the participants to
4 certify that bilinguals could efficiently speak both languages fluently and that monolinguals were
5 not able to communicate in Basque. All the participants acquired their languages before the age of
6 12. Bilinguals and monolinguals didn't differ in any demographic factor (all $ps > .5$), including the
7 age at which they quit formal schooling, the IQ percentile based on the scores they obtained in an
8 abridged version of the Kaufman Brief Intelligence Test, K-BIT (Kaufman & Kaufman, 1990), and
9 the scores in the Spanish version of the Mini Mental State Examination (MMSE, see Lobo et al.,
10 1979). Considering that some researchers have suggested that the level of education can modulate
11 the presence of a bilingual advantage (see Gollan et al., 2011), we recruited participants from all the
12 educational strata but mostly from the lowest one: 9 bilinguals and 10 monolinguals only finished
13 primary school, 2 bilinguals and 3 monolinguals completed middle school, 3 bilinguals and 5
14 monolinguals had a professional training, 6 bilinguals and 4 monolinguals completed high school
15 and 4 bilinguals and 2 monolinguals had completed a university degree. To avoid any cultural
16 difference, they were recruited in the same city (Donostia-San Sebastián) and were non-immigrants.
17 Furthermore, based on self-ratings, they didn't differ significantly in general, speaking and
18 comprehension abilities in Spanish (all $ps > .6$, see Table A.1 for detailed information), which was
19 the language in which they were spoken to and tested in during the whole process.

20

21 -- Table A.1 around here --

22

23 **Materials**

24 **Experiment 1A.** For the *verbal Stroop* task (which was a variation of the classic Stroop task;
25 Stroop, 1935), the names of the colors red, blue and yellow (“rojo”, “azul” and “amarillo” in
26 Spanish) and three pairwise-matched non-color words of a similar length, frequency and syllabic
27 structure (“ropa”, “avión” and “apellido”, translated as clothes, plane and surname, respectively)
28 were arranged to create the Congruent (a color name printed in the same color that it indicates, e.g.,
29 the word “rojo” in red), Incongruent (a color name printed in a different color, e.g., the word “rojo”
30 in blue) and Neutral Word (words that weren't color names printed in the three colors, e.g., the
31 word “ropa” in red) conditions, and a Neutral Symbol (a string of dollar signs printed in the three
32 colors, e.g., “\$\$\$\$” in red) condition was added to have a condition unaffected by language. Each
33 condition consisted in 24 trials, presenting each color the same amount of times, paired equally to

1 every word in each condition. Each participant was presented with a total of 96 experimental trials,
2 and the trial presentation order was randomized across participants.

3
4 **Experiment 1B.** For the *numerical Stroop* task, 48 stimuli were created using eight digits (1, 2,
5 3, 4, 6, 7, 8 and 9). Each digit was presented the same number of times in each condition (4 times)
6 and in total (12 times). Each trial consisted in two digits (e.g., 1-6), one on the left side and another
7 one on the right side of the screen. Three conditions were created: 16 congruent trials (the larger
8 number in magnitude was also the bigger in size, e.g., small 1-big 6), 16 incongruent trials (the
9 larger number in magnitude was the smaller in size, e.g., big 1-small 6) and 16 neutral trials (two
10 same numbers different in size, e.g., big 1-small 1). In all the conditions “left” and “right” responses
11 were equally distributed.

12 13 **Procedures**

14 All the participants were tested in the BCBL facilities in Donostia-San Sebastián. The
15 experiment was run using DMDX (Forster & Forster, 2003) following the same procedure reported
16 by Duñabeitia and colleagues (Duñabeitia et al., 2014).

17 **Experiment 1A.** For the *verbal Stroop* task, verbal responses were collected through
18 Sennheiser PC151 headsets. Research assistants indicated to the participants that they had to name
19 the color of the ink in which the word on the screen was presented, as quickly and as accurately as
20 possible. They completed a short training phase that consisted of four trials, one per condition.
21 Immediately after this, the experiment began. They first saw a fixation mark for 250ms and then the
22 target word appeared on the screen for 3000 ms. All the strings were presented in uppercase Courier
23 New font on a black background, and the colors were set in the RGB-scale values as follows:
24 blue=0,0,255; red=255,0,0; yellow=255,255,0.

25 **Experiment 1B.** For the *numerical Stroop*, all the technical equipment and software used in
26 this experiment were identical to that reported for the verbal Stroop task. In this occasion,
27 participants were instructed to indicate with the keyboard which of the digits in each pair displayed
28 in the screen was bigger in size, by pressing “L” to indicate “right” side and “S” to indicate “left”
29 side of the screen. After instructions, they completed three practice trials (one per condition) and
30 feedback regarding their accuracy was provided. Immediately after the practice trials, the
31 experimental trials were presented in a random order for each participant. First, a fixation mark was
32 presented in the center of the screen for 300ms in order to capture participants’ attention. Next, the
33 visual display was presented until participants had responded to it or for a maximum of 3500ms. All
34 the digits were presented in Courier New black font on a white background, each digit of each pair

1 on one side of the screen. The whole experimental session lasted around 5 minutes. The order of the
2 tasks was counterbalanced.

4 Results

5 **Experiment 1A.** For the *verbal Stroop task*, first latencies were analyzed. Reaction times
6 above and below 2.5 standard deviations from each participant's mean in each condition (< 3.2% of
7 the data) were excluded from the analysis. With the remaining data a general ANOVA was run
8 including the factors Condition (Congruent, Incongruent, Neutral Words and Neutral Symbols) and
9 Group (Monolinguals and Bilinguals). In this general 4x2 ANOVA, we found a significant effect of
10 Condition [$F(3, 136)=66.88, p<.01$] but neither the main effect of Group nor the interaction (all
11 $F_s<1$). All the descriptive values of conditions across groups are detailed in Table A.2.

12
13 - Table A.2 around here -

14
15 To explore all the possible venues in which differences could emerge, we also computed 2x2
16 ANOVAS to analyze the potential differences between the groups in the *classic Stroop effect* (i.e.,
17 Congruent vs. Incongruent conditions), *incongruity effect* (Neutral Word vs. Incongruent
18 conditions) and *congruency effect* (Neutral Word vs. Congruent conditions)¹. We further tested
19 these differences with the Bayesian Null Hypothesis Testing (Rouder et al., 2009; Wetzels et al.,
20 2011), comparing bilinguals to monolinguals for every index (i.e., comparing the indices of *Stroop*,
21 *congruency* and *incongruity* of both groups) testing the H_0 of no differences against the H_1 that says
22 that bilinguals should have smaller indices than monolinguals (i.e., a test of the so-called bilingual
23 advantage) using Bayesian t-tests.

24 *Stroop effect* was significant [$F(1, 46)=114.95, p<.01$] but there was no effect of Group nor
25 an interaction between them (all $F_s<1$), which was also supported by the Bayesian t-test (Bayesian
26 t-test of the index between groups: $BF_{01}>5.55$). The same pattern was obtained for the *incongruity*

¹ In order to explore the possibility that the bilingual advantage may be circumscribed to low-educated bilinguals (cf. Gollan et al., 2011), a reanalysis of the data was done with the subset of participants with the lowest educational level. To this end, the 9 bilinguals and 10 monolinguals that had only completed primary school were selected, and the ANOVAs on the RT data failed to show a significant effect of Group, nor an interaction between the two main factors [$F_s<1$]. The analysis of each index reinforced this result, showing a sizeable *Stroop effect* [$F(1,17)=55.40, p<.01$], a marginal *incongruity effect* [$F(1,17)=4.25, p<.06$], and a significant *congruency effect* [$F(1,17)=46.12, p<.01$], which did not interact with the factor Group [all $F_s<1.55$ and $ps>.23$]. The analysis of the error rates mimicked these same results, with all the main effects of Group or interactions with this factor resulting negligible [all $F_s<1$].

1 effect, with significant main effect [$F(1, 46)= 46.42, p<.01$] with no effect of Group neither an
2 interaction (all $F_s<1$) and no differences between groups when the *incongruity effect* was compared
3 across groups using Bayesian approach (Bayesian t-test of the index between groups: $BF_{01}>4.1$).
4 *Congruency effect* was also significant [$F(1, 46)= 30.29, p<.01$] but neither the main effect of
5 Group nor the interaction was significant (all $F_s<1$), and the null hypothesis was supported by the
6 Bayesian t-test (Bayesian t-test of the index between groups: $BF_{01}>5.04$).

7 The error rate analysis also showed a quite similar pattern, first with a general 4x2 ANOVA
8 and then comparing each index separately by using both 2x2 ANOVAs and Bayesian Null
9 Hypothesis Testing between groups. A general 4x2 ANOVA showed a significant effect of
10 Condition [$F(3, 138)= 8.15, p<.01$] and no main effect of Group nor an interaction ($F_s<1$). In the
11 analysis of the *Stroop index*, we observed a main Condition effect [$F(1, 46)=7.27, p<.02$] but
12 neither the main effect of Group nor the interaction was significant ($F_s<1$), indicating no
13 differences in this index between groups (Bayesian t-test of the index between groups: $BF_{01}>3.48$).
14 When the *congruency effect* was explored, we observed no effect of Condition, Group or an
15 interaction (all $F_s<1$), and Bayesian analysis comparing the index across group indicated that the
16 null was slightly more likely than the alternative hypothesis (Bayesian t-test of the index between
17 groups: $BF_{01}>1.42$). Finally, the analysis of the *incongruity effect* showed a main effect of
18 Condition [$F(1, 46)=87.53, p<.01$] but no main effect of Group or interaction ($F_s<1$), and a Bayes
19 factor analysis that favored the null hypothesis over the alternative (Bayesian t-test of the index
20 between groups: $BF_{01}>3.92$).

21

22

23 **Experiment 1B.** The same rationale as in the previous task was followed in the analysis of the
24 *numerical Stroop* task results. In the latencies, responses above and below 2.5 standard deviations
25 from each participant's mean in each condition (< 2.9% of the data) were excluded from the
26 analysis. After trimming, a general 3 x 2 ANOVA was run with the factors Condition (Congruent,
27 Incongruent and Neutral) and Language (Monolinguals and Bilinguals). Analysis showed only a
28 main effect of Condition [$F(2, 92)= 35.07, p<.01$, all other $F_s<1$]. See Table A.2 for detailed
29 information for each condition.

30 As in the previous task, we also explored the *classic Stroop effect* (i.e., Congruent vs.
31 Incongruent conditions), the *incongruity effect* (Neutral vs. Incongruent conditions) and the

1 *congruency effect* (Neutral vs. Congruent conditions)² using both ANOVAs and Bayesian Null
2 Hypothesis Testing to compare the indices between groups. *Stroop* and *incongruity* effects showed
3 the same pattern, with strong Condition effects (all $ps < .01$) and no other main effect or interaction
4 (all $F_s < 1$), while Bayesian analysis also showed that the Null Hypothesis of no differences was
5 more likely than the alternative one of smaller effects for bilinguals ($BF_{01} > 3.6$ and $BF_{01} > 5.72$ for
6 the Bayesian t-tests of the *Stroop* and *incongruity* effects, respectively). In the analysis of the
7 *congruency* effect we also found a main effect of Condition [$F(1, 46) = 7.25, p < .02$], no effect of
8 Group ($F < 1$) and no interaction [$F(1, 46) = 2.91, p > .1$], but the results coming from the Bayesian
9 Null Hypothesis testing didn't allow us to reach any conclusion (Bayesian t-test of the index
10 between groups: $BF_{01} > 0.58$).

11 In a similar vein, the general 3x2 ANOVA on the error rates indicated that there was a strong
12 effect of Condition [$F(2, 92) = 7.23, p < .01$] but no Group effect nor an interaction ($F_s < 1$). When the
13 indices were explored individually, we observed that the *Stroop effect* was significant [$F(1, 46) =$
14 $8.00, p < .01$] but no main effect of Group nor an interaction was found (all $F_s < 1$), coherent with the
15 results from the Bayesian t-test (Bayesian t-test of the index between groups: $BF_{01} > 3.48$).
16 *Incongruity* effect analysis showed a significant effect of Condition [$F(1, 46) = 6.67, p < .02$] and no
17 effect of Group or interaction (all $F_s < 1$), while Bayesian analysis supported the null hypothesis
18 (Bayesian t-test of the index between groups: $BF_{01} > 3.92$). When the *congruency effect* was
19 analyzed, neither the main effect of Condition [$F(1, 46) = 1.00, p > .32$], nor the main effect of Group
20 [$F(1, 46) = 1.00, p > .32$] neither an interaction between them was found [$F(1, 46) = 1.87, p > .18$]; and
21 Bayesian factor analysis showed that the null hypothesis was slightly more likely than the
22 alternative one (Bayesian t-test of the index between groups: $BF_{01} > 1.42$).

23

24 **Interim conclusion**

25 Results of the two tasks in this first experiment clearly show no differences between the
26 monolingual and the bilingual group in the critical measures of both inhibitory (*Stroop effect*) and
27 monitoring skills (overall reaction times). Importantly, the tasks worked as expected, replicating

² As in Experiment 1A, a reanalysis was run on those participants who only completed the obligatory primary school (9 bilinguals and 10 monolinguals). The ANOVAs on the RT data did not show any significant effect of Group or an interaction between the two main factors [$F_s < 1.22$ and $ps > .3$]. The analysis of each index reinforced this result, showing a significant *Stroop effect* [$F(1, 17) = 10.50, p < .01$], and a significant *incongruity effect* [$F(1, 17) = 9.42, p < .01$]. The *congruency effect* was not significant [$F < 1$], and none of the interactions of these indices with the factor Group resulted significant [all $F_s < 1.06$, all $ps > .31$]. The analysis of the error rates mimicked these same results, with all the main effects of Group or interactions with this factor resulting negligible [all $F_s < 2.45$, and $ps > .14$].

1 previous findings on the main indices in both reaction times and error rates with significant and
2 strong *Stroop effects*, mainly due to an *incongruity effect*. Importantly, when carefully matched
3 monolinguals and bilinguals were compared using the conservative Bayesian approach, results
4 clearly favor the null hypothesis, indicating that the data is much more likely to be explained by the
5 hypothesis of “no differences” than by any other alternative model. Furthermore, the potential
6 impact of the educational level on the emergence of the so-called bilingual advantage was also
7 considered. While most of the participants tested were not highly educated, this was not the case for
8 all of them, and considering that the differential effects could be stronger in low-educated samples
9 (see Gollan et al., 2011), we ran an additional set of analyses including only the participants with
10 the lowest education levels and parallel results were obtained.

11 However, instead of a clear cut dichotomous distinction between monolinguals and bilinguals,
12 it might be the case that L2 proficiency modulates the effect. To investigate this potential
13 modulation, we conducted a second experiment using the same tasks with a different sample: a set
14 of bilingual participants who differ in L2 proficiency, ranking from a very limited knowledge to
15 perfectly fluent and balanced bilinguals.

16 **Experiment 2A and 2B: Effect of L2 proficiency in executive control.**

17 **Method**

18 **Participants**

19 70 Basque-Spanish bilingual seniors (45 females; mean age of 69.36, SD=4.40; age range =
20 61-81) were recruited from the same city from the Basque Country (Donostia-San Sebastián) and
21 were non-immigrants. All of them rated themselves as highly proficient in Spanish (average rating
22 in a 1-to-10 scale was 8.72; SD=1.08) while they showed as a group great variability in their
23 Basque General Proficiency, ranging from 1 (very poor level) to 10 (perfectly fluent), which was
24 also confirmed in the personal interviews guided by bilingual native speakers from the research
25 center. All of them had acquired their two languages before the age of 12 (see Table A.3 for
26 detailed information about participants’ profiles). None of them reported history of chronic
27 neuropsychological disorder(s) and all of them had normal or corrected-to-normal vision. All
28 participants signed the written informed consent form approved by the Ethics and Research
29 Committees of the Basque Center on Cognition, Brain and Language (BCBL), and they completed a
30 cognitive screening that included, as in Experiment 1, the Spanish version of the MMSE (Lobo et
31 al., 1979) and an abridged version of the Kaufman Brief Intelligence Test, K-BIT (Kaufman &
32 Kaufman, 1990). Also, these participants were recruited from every different educational strata: out

1 of the 70 participants, 20 had only completed primary school, 7 had completed middle school, 12
2 had a professional training, 12 had completed high school, and 19 got a university degree.

3

4

-- Table A.3 around here --

5

6 **Materials**

7 Materials used for Experiments 2A and 2B were the exact same ones used in the Experiment
8 1A and 1B respectively.

9 **Procedures**

10 The procedure followed for the Experiments 2A and 2B was the same that we followed in the
11 Experiment 1A and 1B respectively.

12

Results

13 In this experiment we performed the same analysis used in Experiment 1, but instead of a
14 between-subject factor separating bilinguals from monolinguals in two groups, we considered
15 Basque General Proficiency as a covariate. This rating scale varied from 1 (very poor level of
16 Basque) to 10 (very fluent) among our 70 bilingual speakers (see Figures B.1 and B.2).

17 **Experiment 2A.** In the *verbal Stroop task*, reaction times above and below 2.5 standard
18 deviations from each participant's mean in each condition (< 2.9% of the data) were excluded from
19 the analysis.

20 With the remaining latencies, a four way ANOVA (Condition: Congruent, Incongruent,
21 Neutral Symbol and Neutral Words) was conducted, showing a main effect of Condition [$F(3,$
22 $207)=168.69, p<.01$]. Paired t-tests showed that *Stroop* [$t(69)=17.37, p<.01$], *incongruity* [$t(69)=$
23 $10.33, p<.01$] and *congruency* [$t(69)=7.96, p<.01$] effects were significant (see Table A.4 for
24 descriptive values).

25 After this general analysis of our bilingual participants, a four way ANCOVA was run
26 including the factor Condition (Congruent, Incongruent, Neutral Words and Neutral Symbols) and
27 using the Basque General Proficiency as covariate³. A significant effect of Condition [$F(3, 204)=$

³ Considering previous findings that show a relation between the intelligence scores and the different executive functioning tasks such as the Stroop task (Adelman, 2002; Arffa, 2007), we also ran a four way ANCOVA including both Basque General Proficiency and IQ percentile values (obtained from the K-BIT) as covariates. Results show that there is a main effect of Condition [$F(3, 201)=21.02, p<.01$], IQ [$F(1, 67)=10.55, p<.01$] and an interaction between them [$F(3, 201)= 7.29,$

1 18.29, $p < .01$] was found, but the effect of Basque General Proficiency was not significant [$F(1,$
2 68)=2.11, $p > .15$], and it didn't interact with Condition ($F < 1$).

3 We also explored each index separately by two-way ANCOVAs to see if there was any
4 modulation of the covariate in the effect. The *classic Stroop effect* (i.e., Congruent vs. Incongruent
5 conditions) was significant [$F(1, 68)=29.49, p < .01$] but the main effect of Basque General
6 Proficiency was not [$F(1, 68)= 1.05, p > .31$], and Basque General Proficiency did not modulate the
7 main effect of Condition ($F < 1$). The *incongruity effect* (Neutral Word vs. Incongruent Conditions)
8 followed the same pattern, with main effect of Condition [$F(1, 68)=16.74, p < .01$], no effect of
9 Basque General Proficiency [$F(1, 68)=2.06, p > .16$] and no modulation of the Basque General
10 Proficiency in the main effect of Condition ($F < 1$). Finally, the *congruency effect* (Neutral Word vs.
11 Congruent conditions) showed a marginal effect [$F(1, 68)=2.96, p < .1$], with no effect of Basque
12 General Proficiency [$F(1, 68)=1.96, p > .17$] nor an interaction between the two main effects [$F(1,$
13 68)=1.20, $p > .28$].

14 In the error rate analysis, the general four-way ANCOVA showed that none of the effects or
15 interactions were significant [all F s < 2 and p s > .17].

16 Considering that preceding studies have proposed the existence of a close relationship
17 between the educational level of the participants and their performance in Stroop-like tasks (see
18 Moering, Schinka, Mortimer, & Graves, 2004), and taking into account that the so-called bilingual
19 advantage has been claimed to depend on this factor (cf. Gollan et al., 2011), we decided to run an
20 additional analysis in order to shed light on this issue. A four-way ANCOVA was run including the
21 factor Condition (Congruent, Incongruent, Neutral Words and Neutral Symbols) and using the
22 Basque General Proficiency and Education (i.e., the age at which participants quit formal education)
23 as covariates. Results showed a significant main effect of Condition [$F(3,201)=9.06, p < .01$] and a
24 marginal effect of Education [$F(1,67)=3.17, p = .08$], with no effect of Basque General Proficiency
25 [$F(1,67)=1.85, p > .18$], nor an interaction between Condition and any of the covariates [all F s < 1].

26 When each index was explored independently in the corresponding set of two-way ANCOVAs, the

$p < .01$]. When looking at each index, analysis of the *Stroop effect* revealed a main effect of
Condition [$F(1, 67)=48.86, p < .01$], IQ [$F(1, 67)=4.86, p < .01$] and an interaction between them
[$F(1, 67)=15.25, p < .01$] indicating that the *Stroop effect* was smaller for higher IQ values ($r = -0.44,$
 $p < .01, n = 70$). We observe a similar pattern in the *congruency effect*, with a significant Condition
[$F(1, 67)=19.08, p < .01$] and IQ [$F(1, 67)= 7.21, p < .01$] effects as well as an interaction between
them [$F(1, 67)= 18.16, p < .01$] indicating that the *congruency effect* was smaller for higher IQ scores
($r = -.47, p < .01, n = 70$). Finally, *incongruity effect* analysis revealed a main effect of Condition [$F(1,$
67)= 7.27, $p < .01$] and IQ [$F(1, 67)= 15.19, p < .01$] but no other effect was significant (all F s < 1).
Basque General Proficiency was not significant (all F s < 1, all p s > .33) for any of the indices, nor did
it interact with condition.

1 *Stroop effect* (Congruent vs. Incongruent) resulted significant [$F(1,67)=15.45, p<.01$], but it was not
2 modulated by Basque General Proficiency or Education [$F_s<1$]. The main effects of Basque
3 General Proficiency and Education were not significant either [$F_s(1,67)<2.6, ps>.11$]. The
4 *incongruity effect* (Incongruent vs. Neutral Word) was significant [$F(1,67)=8.71, p>.01$], but it was
5 not modulated by Basque General Proficiency or Education [all $F_s<1$]. The main effect of
6 Education was marginal [$F(1,67)=3.02, p=.09$], but the effect of Basque General Proficiency was
7 negligible [$F(1,67)=1.81, p>.18$]. Finally, the *congruency effect* (Congruent vs. Neutral Word) was
8 not significant [$F(1,67)=1.54, p>.22$], and it was not modulated by Basque General Proficiency
9 [$F(1,67)=1.14, p<.3$] or Education [$F<1$]. The main effects of Basque General Proficiency and
10 Education were not significant [$F(1,67)=1.73, p>.19$ and $F(1,67)=2.54, p>.12$, respectively]. The
11 four-way ANCOVA run on the accuracy data showed that none of the effects or interactions was
12 significant (all $F_s<1.8$ and $ps>.2$). Altogether, the marginal main effects of Education that emerged
13 in the general ANCOVA showed that overall reaction times tended to be shorter for people with
14 higher educational level, but critically, this analysis demonstrated that Education did not modulate
15 any of the indices of interest.

16

17 To further check for any possible modulation of the indices as a function of the relevant
18 demographic data collected, different correlation analyses were run between the factors of interest
19 and the indices obtained (both for RTs and error rates). Crucially for the purposes of this study,
20 Basque General Proficiency did not correlate with the *Stroop* ($r=.07, p>.58$), *incongruity* ($r=-.003,$
21 $p>.98$) or *congruency* ($r=.13, p>.28$) indices in the RTs (see Figure B.1). The error rate analysis
22 showed the same pattern, and neither the *Stroop* ($r=.04, p>.74$), nor the *congruency* ($r=-.05, p>.66$)
23 and the *incongruity* ($r=.06, p>.64$) indices were correlated with the general proficiency that
24 participants had in Basque (Footnote 3).

25

26 - Figure B.1 around here -

27

28

29

30 **Experiment 2B.** Reaction times above and below 2.5 standard deviations from each
31 participant's mean in each condition (< 3.2% of the data) of the *numerical Stroop task* were
32 excluded from the analysis. After trimming, a three way ANCOVA was run including the factor

1 Condition (Congruent, Incongruent and Neutral) and Basque General Proficiency as a covariate⁴
2 (see Table A.4 for descriptives). The main effect of Condition was significant [$F(2, 136)= 7.39$,
3 $p<.01$], but Basque General Proficiency was not, and it did not interact with Condition either (all
4 $F_s<1$). Two-way ANCOVAS to account for each effect and its modulation by Basque General
5 Proficiency, if any, showed that both the *Stroop* [$F(1, 68)=12.54, p>.01$] and the *congruency* [$F(1,$
6 $68)=7.81, p<.01$] indices were significant but Basque General Proficiency and its interaction with
7 Condition were not (all $F_s<1$). The *incongruity* effect was not statistically significant [$F(1,$
8 $68)=2.93, p<.1$], and Basque General Proficiency was not and it did not interact with the Condition
9 effect ($F_s<1$).

10
11 - Table A.4 around here -
12

13 The general three-way ANCOVA run in error rates showed that neither Condition [$F(2,$
14 $136)=.17, p>.84$] nor Basque General Proficiency nor their interaction were significant ($F_s<1$).

15 As in the Experiment 1, the potential effect of Education (i.e., the age at which participants
16 quit formal education) was investigated in a three-way ANCOVA including Basque General
17 Proficiency and Education as covariates. The main effect of Condition [$F(2,134)=5.40, p<.01$] and
18 Education [$F(1,64)=4.61, p<.04$] resulted significant. Crucially, the effect of Condition did not
19 interact with any of the other factors [all $F_s<1.5$, all $ps>.22$], and the effect of Basque General
20 Proficiency was not significant [$F<1$]. The analysis of the *Stroop effect* showed significant effects
21 of Condition [$F(1,67)=7.78, p<.01$] and Education [$F(1,67)=5.22, p<.03$], with no other relevant
22 effects or interactions [all $F_s<1$]. Similarly, the analysis of the *incongruity effect* showed a main
23 effect of Condition [$F(1,67)= 4.86, p<.04$] and of Education [$F(1,67)=4.52, p<.04$], with no main
24 effect of Basque General Proficiency [$F<1$] or interaction between Condition and Basque General

⁴ As in Experiment 2A, we ran a general ANCOVA including both Basque General Proficiency and IQ percentile values as covariates. Condition [$F(2, 134)=12.37, p<.01$], IQ [$F(1, 67)=14.54, p<.01$] and the interaction between them [$F(2, 134)= 6.08, p<.01$] were significant. Exploring each index, the analysis of the *Stroop effect* revealed a main effect of Condition [$F(1, 67)=18.43, p<.01$], IQ [$F(1, 67)=16.50, p<.01$] and an interaction [$F(1, 67)=5.78, p<.02$], showing that the *Stroop effect* decreased as IQ values increased ($r= -0.28, p<.02, n= 70$). In the *congruency effect* we found a significant IQ effect [$F(1, 67)= 10.89, p<.01$], with no other significant main effects or interactions (all $F_s<2$, all $ps>.21$). The *incongruity effect* analysis revealed a main effect of Condition [$F(1, 67)= 10.71, p<.01$] and IQ [$F(1, 67)= 15.53, p<.01$], as well as an interaction [$F(1, 67)=8.04, p<.01$], indicating reduced *incongruity effects* for higher IQ values ($r= -.34, p<.01, n=70$). Basque General Proficiency was not significant in any of the analyses, nor did it interact with Condition (all $F_s<1$ and $ps>.53$).

1 Proficiency [$F < 1$], neither between Condition and Education [$F(1,67) = 2.16, p < .15$]. The analysis of
2 the *congruency effect* showed no main effect or interactions [all F s < 1.5 and p s $> .24$], except for a
3 marginal effect of Education [$F(1,67) = 3.8, p < .06$]. The main effects of Education that consistently
4 emerged in these analyses showed that participants who quit formal education later were the ones
5 associated with faster reaction times in the different conditions. The three-way ANCOVA on the
6 accuracy data showed that none of the main effects or interactions were significant [all F s < 1].

7 Additionally, the possible relationship between the demographic variables of interest and
8 the indices measured in this Numerical Stroop task was explored in a correlation analysis. Crucially
9 for the hypothesis explored in this study, we observed that the *Stroop* ($r = .03, p > .8$), *congruency* ($r =$
10 $.1, p > .41$) and *incongruity* ($r = .08, p > .5$) indices were not correlated with the General Basque
11 Proficiency in the RT analysis (see Figure B.2). A similar pattern was observed for the error rates,
12 with none of the indices being correlated with the general proficiency that participants had in
13 Basque (*Stroop*: $r = .13, p > .30$; *congruency*: $r = .10, p > .40$; *incongruity*: $r = .09, p > .44$; Footnote 4).

14
15 - Figure B.2 around here -
16

17 **Interim Conclusion**

18 In this experiment we obtained significant and strong Stroop effects in the latency analysis of
19 both tasks, mainly due to the incongruity effects. However, when the impact of Basque General
20 Proficiency in the different indices was analyzed, and even when the variance provided by factors
21 such as IQ or Educational Level was controlled for, the ANCOVAs showed no significant effect of
22 participants' knowledge of a second language neither a modulation of the main indices based on the
23 this knowledge, as measured by their Basque General Proficiency.

24 **General Discussion**

25 In the current study, we analyzed the effects derived from lifelong bilingualism in domain-
26 general cognitive abilities related to inhibitory control and monitoring skills in samples of elderly
27 bilinguals and monolinguals. In Experiments 1A and 1B, we compared monolingual Spanish-
28 speaking old adults and elderly Spanish-Basque bilinguals who have been immersed in a bilingual
29 society and who have used their two languages during the vast majority of their lives on a daily
30 basis. All of them were non-immigrants coming from the same city and did not differ in any of the
31 demographic factors nor in the linguistic skills in Spanish (the language that both groups shared and
32 the language in which they were tested). These participants were presumably in a declining process
33 of their cognitive abilities due to normal aging, although their cognitive functioning was at normal
34 levels according to the scores obtained in the Mini Mental State Examination (MMSE, see Folstein,

1 Folstein & McHugh, 1975; i.e., all participants scored above 26 in Experiment 1 (Median = 29.5),
2 and above 24 in Experiment 2 (Median = 29). These bilinguals were tentatively selected as a good
3 test case to explore any enhancing effect that bilingualism may have in the inhibitory control and
4 monitoring skills, because they have been exposed to lifelong bilingualism and they are not at their
5 peak cognitive abilities. The results unambiguously demonstrated a complete absence of differences
6 between lifelong bilingual seniors and their monolingual peers either in monitoring abilities (which
7 would have been reflected in overall faster reaction times) or in inhibitory skills (which would have
8 been shown by reduced conflict or Stroop effects). Importantly, when the same hypotheses were
9 tested only in the subsets of seniors with the lowest educational levels following Gollan et al.'s
10 (2011) rationale, the same results were replicated, demonstrating that the lack of a bilingual
11 advantage does not circumscribe to certain levels of education.

12 In Experiments 2A and 2B, a different approach was taken in order to delve into this same
13 issue, this time testing a large group of bilinguals that differed in their L2 proficiency, ranging from
14 a clearly low knowledge of the second language to perfectly fluent and balanced bilinguals. If
15 bilingualism has any impact in cognitive functioning, then this impact should be modulated by the
16 degree of knowledge of the second language, and thus the effects obtained in both tasks should
17 show a significant correlation with seniors' L2 proficiency. Preceding research on this issue has
18 failed at providing a consistent picture. Thus, to the best of our knowledge, the current study is the
19 first one that aimed at checking for any possible modulation of bilingual seniors' inhibitory
20 capacities by their degree of mastery of an L2 along a continuum, keeping all the other factors
21 static. The 70 seniors tested in the second group of experiments came from the same city and all of
22 them had acquired their second language before the age of 12, meaning that, despite individual
23 differences in their use of the languages, the general degree of exposure to the languages in social
24 contexts could be considered as highly homogeneous. The results from the ANCOVAs and
25 correlations demonstrated that, regardless of their L2 proficiency participants showed comparable
26 inhibitory skills (as measured by the Stroop effects), thus extending and qualifying earlier evidence
27 obtained from younger bilinguals who did not show a significant relationship between L2
28 proficiency and the size of inhibitory control measures.

29
30 When the ANCOVAs were run including the IQ percentiles obtained from the K-BIT as a
31 covariate, we replicated previous findings showing that the Stroop indices are reduced for higher IQ
32 values (see Adelman, 2002; Arffa, 2007), but still we observed no main effect of Basque General
33 Proficiency, nor an interaction of it with any of the indices. Similarly, when additional ANCOVAs
34 were run including Education as a covariate, these same results were replicated. Participants'

1 performance in the Stroop task improved as a function of Education (see also Houx, Jolles, &
2 Vreeling, 1993; Moering, Schinka, Mortimer, & Graves, 2004; Van der Elst, Van Boxtel, Van
3 Breukelen, & Jolles, 2006), but no effect of Basque General Proficiency was found, nor an
4 interaction between Basque General Proficiency and any of the indices of interest.

5
6 Hence, the current results demonstrate that when the hypothesis of a potential modulation of
7 inhibitory skills as a function of increased L2 proficiency is tested with lifelong elderly, no
8 significant difference in their executive functions was observed as a function of their L2 knowledge.

9 The results found here add to the growing body of evidence that is gaining strength in the last
10 several years and that suggest that the bilingual advantage in executive functioning (and explicitly
11 in inhibitory and monitoring abilities) is actually non-existent. Different arguments have been used
12 to reconcile the failures to obtain bilingual advantages. The absence of evidence favoring the
13 bilingual advantage in young adults (Bialystok, Martin, & Viswanathan, 2005) has been argued to
14 be a consequence of the ceiling effects that cognitive skills might feature in those ages; and
15 therefore it might be captured easier in children. However, this has not been the case when
16 contrasting a large sample of monolingual and bilingual children (Duñabeitia et al., 2014; Antón et
17 al., 2014, Gathercole et al., 2014). In turn, this absence of evidence of an enhancement of general
18 executive functioning in young children could also be argued to be due to the lack of enough
19 exposure to bilingualism, meaning that these bilinguals have not undergone sufficient training in
20 their lives. Consequently, it has been argued that the so-called bilingual advantage might emerge in
21 later stages of life, given that the benefits of lifelong bilingualism could be better observed in
22 samples of seniors whose cognitive skills are presumably declining. Nonetheless, here we
23 demonstrated that in our elderly bilingual sample the bilingual advantage is absent. Furthermore, it
24 is worth noting that our data is not the only one showing unambiguously that there are no
25 differences in executive functioning between elderly monolinguals and bilinguals (Kirk et al., 2014;
26 Kousaie & Phillips, 2012; de Bruin, Bak & Della Sala, 2015).

27 However, we acknowledge that the present results should be considered with cautiousness,
28 and although our data show that bilingualism does not enhance executive functioning in the elderly,
29 benefits derived from bilingualism in other domains should not be overlooked or disregarded, like
30 the obvious benefits in terms of social and communicational skills. Nowadays, other potential
31 paybacks of bilingualism at non-linguistic levels are also under debate, such as its neuro-protective
32 value regarding the delay in the emergence of the symptoms of certain types of dementias (see
33 Bialystok, Craik, & Freedman, 2007; Albán-González & Ortega-Campoverde, 2014). While some
34 researchers do not support this potential consequence of bilingualism when the characteristics of the

1 samples are carefully controlled for (see Chertkow et al., 2010; Lawton, Gasquoine, & Weimer,
2 2015), others report significant results even in carefully matched groups (see Alladi et al., 2013;
3 Woumans et al., 2015). However, as correctly stated by Paap, Johnson and Sawi (2015b), the most
4 compelling pieces of evidence at this regard may come from longitudinal studies following cohorts
5 of individuals, and most of these studies yielded non-significant differences, or even monolingual-
6 favoring trends (e.g., Crane et al., 2009; Lawton, Gasquoine, & Weimer, 2015; Sanders, Hall, Katz,
7 & Lipton, 2012; Yeung, St. John, Menec, & Tyas, 2014; Zahodne, Schofield, Farrell, Stern, &
8 Manly, 2014), while only one presented evidence in favor of a bilingual advantage at this level of
9 analysis (Wilson et al., 2014).

10 The vast majority of research exploring the so-called bilingual advantage focuses on group
11 comparisons, but conclusions from this type of experimental designs are always to be taken with
12 caution, because despite all the effort put in matching samples, one might consider that dozens of
13 factors can still play a role when the main comparison of interest is done based on a non-randomly
14 distributed variable, i.e., bilingualism. In our present study investigating Basque-Spanish bilinguals
15 and Spanish monolinguals, one could argue that Basque speakers represent a cultural minority with
16 a different social and historical background from those of Spanish monolinguals. However, it is
17 important to mention at this regard that our participants come from the same city and that they are
18 not immigrants, thus sharing cultural and historical background to a great extent, and being as
19 comparable as a between-subjects design allows for (Experiment 1). Furthermore, in an attempt to
20 reduce the potential impact of uncontrolled between-groups factors, a second series of experiments
21 were run considering the proficiency in the second language as a continuum in a sample of bilingual
22 individuals (Experiment 2). Nonetheless, we acknowledge that other alternative approaches should
23 be also favored. As suggested by Paap, Johnson and Sawi (2015b) and Duñabeitia and Carreiras
24 (2015), it would be worth exploring how a wide range of cognitive skills (including executive
25 functioning) changes before and after the acquisition of a language in the same group of individuals,
26 following a longitudinal approach. According to recent evidence, it could be tentatively predicted
27 that in case of the emergence of a difference (i.e., a bilingual advantage), this would be most clearly
28 seen during the first years of immersion in an L2 context (see Heidlmayr, Moutier, Hemforth,
29 Courtin, Tanzmeister, & Isel, 2014).

30 **Conclusions**

31 If we consider the evidence presented here together with the other published results showing
32 no bilingual advantage in young children when critical confounding factors are controlled for, the
33 argument that the ceiling effect of cognitive abilities is responsible for the lack of bilingual

1 advantage in young adulthood is weakened. If that was the case, the advantage would be elusive in
2 young adulthood but strong in children and the elderly. In addition, the present data (and in other
3 studies cited above) showing no benefits of lifelong bilingualism weakens the hypothesis that there
4 were no evident differences in children because lifelong exposure to bilingualism is needed to
5 achieve that boosting effect in executive functioning. Thus, the evidence we collected in the present
6 study adds to the growing body of evidence showing a comprehensive picture indicating that a
7 bilingual advantage in tasks measuring executive control in any segment of the population is very
8 likely to be produced by uncontrolled non-linguistic factors, rather than by the critical between-
9 group difference of being bilingual or monolingual. As recently suggested, when those factors are
10 controlled for and participant groups are carefully matched, no significant differences are captured
11 between monolinguals and bilinguals in their performance in tasks tapping into inhibition and
12 monitoring abilities (Paap, Johnson & Sawi, 2015a, 2015b). In fact, we failed to find any signs of a
13 bilingual advantage in samples of carefully selected lifelong non-immigrant native bilingual seniors
14 with different degrees of proficiency that live in a bilingual society, when their performance is
15 compared with that of a group of carefully matched monolingual seniors. Hence, we conclude that
16 lifelong bilingualism does not represent any specific benefit in executive functions in healthy
17 elderly (see also Hilchey, Saint-Aubin, & Klein, 2015).

18
19

1 **Acknowledgments:**

- 2 This research has been partially funded by grants PSI2015-65689-P and SEV-2015-0490 from the
3 Spanish Government, PI2015-1-27 from the Basque Government, AThEME-613465 from the
4 European Union and an ERC-AdG-295362 grant from the European Research Council.

References:

- 1
2 Adelman, N., Menon, V., Blasey, C. M., White, C. D., Warsofsky, I. S., Glover, G. H., & Reiss, A.
3 L. (2002). A developmental fMRI study of the Stroop color-word task. *NeuroImage*, *16*,
4 61– 75. <http://dx.doi.org/10.1006/nimg.2001.1046>
- 5 Albán-González, G., & Ortega-Campoverde, T. (2014). Relationship between bilingualism and
6 Alzheimer's. *Suma de Negocios*, *5*(11), 126-133. [http://dx.doi.org/10.1016/S2215-](http://dx.doi.org/10.1016/S2215-910X(14)70027-8)
7 [910X\(14\)70027-8](http://dx.doi.org/10.1016/S2215-910X(14)70027-8)
- 8 Alladi, S., Bak, T. H., Duggirala, V., Surampudi, B., Shailaja, M., Shukla, A. K., Chaudhuri, D.
9 M., & Kaul, S. (2013). Bilingualism delays age at onset of dementia, independent of
10 education and immigration status. *Neurology*, *8*, 1938-1944.
11 <http://dx.doi.org/10.1212/01.wnl.0000436620.33155.a4>
- 12 Antón, E., Duñabeitia, J.A., Estévez, A., Hernández, J.A., Castillo, A., Fuentes, L.J., Davidson, D.
13 J., & Carreiras, M., (2014). Is there a bilingual advantage in the ANT task? Evidence from
14 children. *Frontiers in Psychology*, *5*, 398, 1-12. <http://dx.doi.org/10.3389/fpsyg.2014.00398>
- 15 Arffa, S. (2007). The relationship of intelligence to executive function and non-executive function
16 measures in a sample of average, above average, and gifted youth. *Archives of Clinical*
17 *Neuropsychology*, *22*(8), 969-978. <http://dx.doi.org/10.1016/j.acn.2007.08.001>
- 18 Besner, D., & Coltheart, M. (1979). Ideographic and alphabetic processing in skilled reading of
19 English. *Neuropsychologia*, *17*, 467–472. [http://dx.doi.org/10.1016/0028-3932\(79\)90053-8](http://dx.doi.org/10.1016/0028-3932(79)90053-8)
- 20 Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon
21 task. *Canadian Journal of Experimental Psychology*, *60*(1), 68-79.
22 <http://dx.doi.org/10.1037/cjep2006008>
- 23 Bialystok, E., Abutalebi, J., Bak, T. H., Burke, D. M., & Kroll, J. F. (2016). Aging in two
24 languages: Implications for public health. *Ageing Research Reviews*, *27*, 56-60.
25 <http://dx.doi.org/10.1016/j.arr.2016.03.003>
- 26 Bialystok, E., Craik, F. I., & Freedman, M. (2007). Bilingualism as a protection against the onset of
27 symptoms of dementia. *Neuropsychologia*, *45*(2), 459-464.
28 <http://dx.doi.org/10.1016/j.neuropsychologia.2006.10.009>

- 1 Bialystok, E., Craik, F. I., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological*
2 *Science in the Public Interest*, 10(3), 89-129. <http://dx.doi.org/10.1177/1529100610387084>
- 3 Bialystok, E., Craik, F. I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive
4 control: evidence from the Simon task. *Psychology and aging*, 19(2), 290-303.
5 <http://dx.doi.org/10.1037/0882-7974.19.2.290>
- 6 Bialystok, E., Craik, F.I., & Luk, G. (2008). Cognitive control and lexical access in younger and
7 older bilinguals. *Journal of Experimental Psychology: Learning, memory, and*
8 *cognition*, 34(4), 859-873. <http://dx.doi.org/10.1037/0278-7393.34.4.859>
- 9 Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: Consequences for Mind and
10 Brain. *Trends in Cognitive Sciences*, 16(4), 240–250.
11 <http://dx.doi.org/10.1016/j.tics.2012.03.001>
- 12 Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence
13 from the dimensional change card sort task. *Developmental science*, 7(3), 325-339.
14 <http://dx.doi.org/10.1111/j.1467-7687.2004.00351.x>
- 15 Bialystok, E., Martin, M. M., & Viswanathan, M. (2005). Bilingualism across the lifespan: The rise
16 and fall of inhibitory control. *International Journal of Bilingualism*, 9(1), 103-119.
17 <http://dx.doi.org/10.1177/13670069050090010701>
- 18 Bialystok, E., & Shapero, D. (2005). Ambiguous benefits: The effect of bilingualism on reversing
19 ambiguous figures. *Developmental Science*, 8(6), 595-604. [http://dx.doi.org/10.1111/j.1467-](http://dx.doi.org/10.1111/j.1467-7687.2005.00451.x)
20 [7687.2005.00451.x](http://dx.doi.org/10.1111/j.1467-7687.2005.00451.x)
- 21 Chertkow, H., Whitehead, V., Phillips, N., Wolfson, C., Atherton, J., & Bergman, H. (2010).
22 Multilingualism (but not always bilingualism) delays the onset of Alzheimer disease:
23 evidence from a bilingual community. *Alzheimer Disease & Associated Disorders*, 24(2),
24 118-125. <http://dx.doi.org/10.1097/WAD.0b013e3181ca1221>
- 25 Clark, J.L.D. (1981). Language. In T. S. Barrows et al. (Eds.), *College students' knowledge and*
26 *beliefs: A survey of global understanding* (pp. 25-35). New Rochelle, NY: Change
27 Magazine Press.

- 1 Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual
2 advantage in conflict processing: Now you see it, now you don't. *Cognition*, *113*(2), 135-
3 149. <http://dx.doi.org/10.1016/j.cognition.2009.08.001>
- 4 Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution:
5 Evidence from the ANT task. *Cognition*, *106*(1), 59-86.
6 <http://dx.doi.org/10.1016/j.cognition.2006.12.013>
- 7 Costa, A., & Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the
8 brain? *Nature Reviews Neuroscience*, *15*(5), 336-345. <http://dx.doi.org/10.1038/nrn3709>
- 9 Crane, P. K., Gibbons, L. E., Arani, K., Nguyen, V., Rhoads, K., McCurry, S. M., ... & White, L.
10 (2009). Midlife use of written Japanese and protection from late life
11 dementia. *Epidemiology*, *20*(5), 766. <http://dx.doi.org/10.1097/EDE.0b013e3181b09332>
- 12 de Abreu, P. M. E., Cruz-Santos, A., Tourinho, C. J., Martin, R., & Bialystok, E. (2012).
13 Bilingualism enriches the poor enhanced cognitive control in low-income minority
14 children. *Psychological science*, *23*(11), 1364-1371.
15 <http://dx.doi.org/10.1177/0956797612443836>
- 16 de Bruin, A., Bak, T. H., & Della Sala, S. (2015). Examining the effects of active versus inactive
17 bilingualism on executive control in a carefully matched non-immigrant sample. *Journal of*
18 *Memory and Language*, *85*, 15-26. <http://dx.doi.org/10.1016/j.jml.2015.07.001>
- 19 Duñabeitia, J.A., Hernández, J.A., Antón, E., Macizo, P., Estévez, A., Fuentes, L.J., & Carreiras, M.
20 (2014). The inhibitory advantage in bilingual children revisited: myth or reality?
21 *Experimental Psychology*, *61* (3), 234-251. <http://dx.doi.org/10.1027/1618-3169/a000243>
- 22 Duñabeitia, J. A., & Carreiras, M. (2015). The bilingual advantage: Acta est fabula. *Cortex*, *73*,
23 371-372. <http://dx.doi.org/10.1016/j.cortex.2015.06.009>
- 24 Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target
25 letter in a nonsearch task. *Perception & psychophysics*, *16*(1), 143-149.
26 <http://dx.doi.org/10.3758/BF03203267>
- 27 Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": a practical method
28 for grading the cognitive state of patients for the clinician. *Journal of psychiatric*
29 *research*, *12*(3), 189-198. [http://dx.doi.org/10.1016/0022-3956\(75\)90026-6](http://dx.doi.org/10.1016/0022-3956(75)90026-6)

- 1 Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond
2 accuracy. *Behavior Research Methods, Instruments, & Computers*, 35, 116-124.
3 <http://dx.doi.org/10.3758/BF03195503>
- 4 Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Guasch, N. V., ... & Jones,
5 L. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan
6 evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic
7 tasks. *Frontiers in psychology*, 5, 11. <http://dx.doi.org/10.3389/fpsyg.2014.00011>
- 8 Gold, B. T., Kim, C., Johnson, N. F., Kryscio, R. J., & Smith, C. D. (2013). Lifelong bilingualism
9 maintains neural efficiency for cognitive control in aging. *The Journal of*
10 *Neuroscience*, 33(2), 387-396. <http://dx.doi.org/10.1523/JNEUROSCI.3837-12.2013>
- 11 Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects
12 picture naming but not picture classification. *Memory & Cognition*, 33(7), 1220-1234.
13 <http://dx.doi.org/10.3758/BF03193224>
- 14 Gollan, T. H., Salmon, D. P., Montoya, R. I., & Galasko, D. R. (2011). Degree of bilingualism
15 predicts age of diagnosis of Alzheimer's disease in low-education but not in highly educated
16 Hispanics. *Neuropsychologia*, 49(14), 3826-3830.
17 <http://dx.doi.org/10.1016/j.neuropsychologia.2011.09.041>
- 18 Hartshorne, J. K., & Germine, L. T. (2015). When does cognitive functioning peak? The
19 asynchronous rise and fall of different cognitive abilities across the life span. *Psychological*
20 *science*, 26(4), 433-443. <http://dx.doi.org/10.1177/0956797614567339>
- 21 Heidlmayr, K., Moutier, S., Hemforth, B., Courtin, C., Tanzmeister, R., & Isel, F. (2014).
22 Successive bilingualism and executive functions: The effect of second language use on
23 inhibitory control in a behavioural Stroop Colour Word task. *Bilingualism: Language and*
24 *Cognition*, 17(03), 630-645. <http://dx.doi.org/10.1017/S1366728913000539>
- 25 Heilenman, L. K. (1990). Self-assessment of second language ability: The role of response
26 effects. *Language Testing*, 7(2), 174-201. <http://dx.doi.org/10.1177/026553229000700204>
- 27 Hernández, M., Costa, A., Fuentes, L. J., Vivas, A. B., & Sebastian-Galles, N. (2010). The impact
28 of bilingualism on the executive control and orienting networks of attention. *Bilingualism:*
29 *Language and Cognition*, 13(03), 315-325. <http://dx.doi.org/10.1017/S1366728909990010>

- 1 Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference
2 tasks? Implications for the plasticity of executive control processes. *Psychonomic bulletin*
3 *& review*, 18(4), 625-658. <http://dx.doi.org/10.3758/s13423-011-0116-7>
- 4 Hilchey, M. D., Saint-Aubin, J., & Klein, R. M. (2015). Does bilingual exercise enhance cognitive
5 fitness in nonlinguistic executive processing tasks. *Cambridge Handbook of Bilingual*
6 *Processing*. Cambridge: CUP.
- 7 Houx, P. J., Jolles, J., & Vreeling, F. W. (1993). Stroop interference: aging effects assessed with the
8 Stroop Color-Word Test. *Experimental aging research*, 19(3), 209-224.
9 <http://dx.doi.org/10.1080/03610739308253934>
- 10 Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech
11 production?. *Acta psychologica*, 127(2), 277-288.
12 <http://dx.doi.org/10.1016/j.actpsy.2007.06.003>
- 13 Kaufman, A. S., & Kaufman, N. L. (1990). *Manual for the Kaufman Brief Intelligence Test*. Circle
14 Pines, MN: American Guidance Service.
- 15 Kaufmann, L., Koppelstaetter, F., Delazer, M., Siedentopf, C., Rhomberg, P., Golaszewski, S.,
16 Felber, S., & Ischebeck, A. (2005). Neural correlates of distance and congruity effects in a
17 numerical Stroop task: an event-related fMRI study. *Neuroimage*, 25(3), 888-898.
18 <http://dx.doi.org/10.1080/20445911.2014.929580>
- 19 Kirk, N. W., Fiala, L., Scott-Brown, K. C., & Kempe, V. (2014). No evidence for reduced Simon
20 cost in elderly bilinguals and bidialectals. *Journal of Cognitive Psychology*, 26(6), 640-
21 648. <http://dx.doi.org/10.1080/20445911.2014.929580>
- 22 Kousaie, S., & Phillips, N. A. (2012). Ageing and bilingualism: Absence of a “bilingual advantage”
23 in Stroop interference in a nonimmigrant sample. *The Quarterly Journal of Experimental*
24 *Psychology*, 65(2), 356-369. <http://dx.doi.org/10.1080/17470218.2011.604788>
- 25 Kroll, J. F., & Bialystok, E. (2013). Understanding the Consequences of Bilingualism for Language
26 Processing and Cognition. *Journal of Cognitive Psychology*, 25(5).
27 <http://dx.doi.org/10.1080/20445911.2013.799170>

- 1 Lagrou, E., Hartsuiker, R. J., & Duyck, W. (2013). The influence of sentence context and accented
2 speech on lexical access in second-language auditory word recognition. *Bilingualism:
3 Language and Cognition, 16*(03), 508-517. <http://dx.doi.org/10.1017/S1366728912000508>
- 4 Lawton, D. M., Gasquoine, P. G., & Weimer, A. A. (2015). Age of dementia diagnosis in
5 community dwelling bilingual and monolingual Hispanic Americans. *Cortex, 66*, 141-145.
6 <http://dx.doi.org/10.1016/j.cortex.2014.11.017>
- 7 LeBlanc, R., & Painchaud, G. (1985). Self-assessment as a second language placement instrument.
8 *TESOL Quarterly, 19*, 673-687. <http://dx.doi.org/10.2307/3586670>
- 9 Lobo, A., Ezquerra, J., Gómez, F., Sala, J. y Seva, A. (1979) El mini-examen cognoscitivo. Un test
10 sencillo y práctico para detectar alteraciones intelectuales en pacientes médicos. *Actas Luso
11 Españolas de Neurología y Psiquiatría, 7*, 189-201.
- 12 Martin-Rhee, M. M., & Bialystok, E. (2008). The development of two types of inhibitory control in
13 monolingual and bilingual children. *Bilingualism: language and cognition, 11*(01), 81-93.
14 <http://dx.doi.org/10.1017/S1366728907003227>
- 15 Mezzacappa, E. (2004). Alerting, orienting, and executive attention: Developmental properties and
16 sociodemographic correlates in an epidemiological sample of young, urban children. *Child
17 development, 75*(5), 1373-1386. <http://dx.doi.org/10.1111/j.1467-8624.2004.00746.x>
- 18 Midgley, K. J., Holcomb, P. J., Van Heuven, W. J. B., & Grainger, J. (2008). An
19 electrophysiological investigation of cross-language effects of orthographic neighborhood.
20 *Brain Research, 1246*, 123–135. <http://dx.doi.org/10.1016/j.brainres.2008.09.078>
- 21 Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in
22 executive functions four general conclusions. *Current directions in psychological
23 science, 21*(1), 8-14. <http://dx.doi.org/10.1177/0963721411429458>
- 24 Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000).
25 The unity and diversity of executive functions and their contributions to complex “frontal
26 lobe” tasks: A latent variable analysis. *Cognitive psychology, 41*(1), 49-100.
27 <http://dx.doi.org/10.1006/cogp.1999.0734>

- 1 Moering, R.G., Schinka, J.A., Mortimer, J.A., & Graves, A.B. (2004). Normative data for elderly
2 African Americans for the Stroop Color and Word Test. *Archives of Clinical*
3 *Neuropsychology*, 19, 61–71. [http://dx.doi.org/10.1016/S0887-6177\(02\)00219-6](http://dx.doi.org/10.1016/S0887-6177(02)00219-6)
- 4 Morton, J. B., & Harper, S. N. (2007). What did Simon say? Revisiting the bilingual
5 advantage. *Developmental science*, 10(6), 719-726. <http://dx.doi.org/10.1111/j.1467-7687.2007.00623.x>
- 6
- 7 Noble, K. G., Norman, M. F., & Farah, M. J. (2005). Neurocognitive correlates of socioeconomic
8 status in kindergarten children. *Developmental science*, 8(1), 74-87.
9 <http://dx.doi.org/10.1111/j.1467-7687.2005.00394.x>
- 10 Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual 5 advantage in
11 executive processing. *Cognitive Psychology*, 66(2), 232–258.
12 <http://dx.doi.org/10.1016/j.cogpsych.2012.12.002>
- 13 Paap, K. R., Johnson, H. A., & Sawi, O. (2014). Are bilingual advantages dependent upon specific
14 tasks or specific bilingual experiences? *Journal of Cognitive Psychology*, 26(6), 615-639.
15 <http://dx.doi.org/10.1080/20445911.2014.944914>
- 16 Paap, K. R., Johnson, H. A., & Sawi, O. (2015a). Bilingual advantages in executive functioning
17 either do not exist or are restricted to very specific and undetermined circumstances.
18 *Cortex*, 69, 265-278. <http://dx.doi.org/10.1016/j.cortex.2015.04.014>
- 19 Paap, K.R., Johnson, H.A., & Sawi, O. (2015b). Should the search for bilingual advantages in
20 executive functioning continue? *Cortex*, 74, 305-314.
21 <http://dx.doi.org/10.1016/j.cortex.2015.09.010>
- 22 Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for
23 accepting and rejecting the null hypothesis. *Psychonomic bulletin & review*, 16(2), 225-237.
24 <http://dx.doi.org/10.3758/PBR.16.2.225>
- 25 Sanders, A. E., Hall, C. B., Katz, M. J., & Lipton, R. B. (2012). Non-native language use and risk of
26 incident dementia in the elderly. *Journal of Alzheimer's Disease*, 29(1), 99.
27 doi: 10.3233/JAD-2011-111631
- 28 Santens, S., & Verguts, T. (2011). The size congruity effect: Is bigger always
29 more? *Cognition*, 118(1), 94-110. <http://dx.doi.org/10.1016/j.cognition.2010.10.014>

- 1 Simon, J. R., & Rudell, A. P. (1967). Auditory SR compatibility: the effect of an irrelevant cue on
2 information processing. *Journal of applied psychology*, 51(3), 300-304.
3 <http://dx.doi.org/10.1037/h0020586>
- 4 Singh, N., & Mishra, R. K. (2013). Second language proficiency modulates conflict-monitoring in
5 an oculomotor Stroop task: evidence from Hindi-English bilinguals. *Frontiers in*
6 *psychology*, 4, 322. <http://dx.doi.org/10.3389/fpsyg.2013.00322>
- 7 Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental*
8 *psychology*, 18(6), 643-662. <http://dx.doi.org/10.1037/h0054651>
- 9 Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-
10 language comprehension. *Proceedings of the National Academy of Sciences*, 104(30),
11 12530-12535. <http://dx.doi.org/10.1073/pnas.0609927104>
- 12 Van der Elst, W., Van Boxtel, M.P.J., Van Breukelen, G.J.P., & Jolles, J. (2006). The Stroop Color-
13 Word Test: Influence of age, sex, and education; and normative data for a large sample
14 across the adult age range. *Assessment*, 13, 62-79.
15 <http://dx.doi.org/10.1177/1073191105283427>
- 16 Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E. J. (2011).
17 Statistical evidence in experimental psychology an empirical comparison using 855 t
18 tests. *Perspectives on Psychological Science*, 6(3), 291-298.
19 <http://dx.doi.org/10.1177/1745691611406923>
- 20 Woumans, E., Santens, P., Sieben, A., Versijpt, J., Stevens, M., & Duyck, W. (2015). Bilingualism
21 delays clinical manifestation of Alzheimer's disease. *Bilingualism: Language and*
22 *Cognition*, 18 (3), 568-574. <http://dx.doi.org/10.1017/S136672891400087X>
- 23 Yang, S., Yang, H., & Lust, B. (2011). Early childhood bilingualism leads to advances in executive
24 attention: Dissociating culture and language. *Bilingualism: Language and*
25 *Cognition*, 14(03), 412-422. <http://dx.doi.org/10.1017/S1366728910000611>
- 26 Yeung, C. M., St John, P. D., Menec, V., & Tyas, S. L. (2014). Is bilingualism associated with a
27 lower risk of dementia in community-living older adults? Cross-sectional and prospective
28 analyses. *Alzheimer disease and associated disorders*, 28(4), 326.
29 <http://dx.doi.org/10.1097/WAD.0000000000000019>

1 Zahodne, L. B., Schofield, P. W., Farrell, M. T., Stern, Y., & Manly, J. J. (2014). Bilingualism does
2 not alter cognitive decline or dementia risk among Spanish-speaking
3 immigrants. *Neuropsychology*, 28(2), 238. <http://dx.doi.org/10.1037/neu0000014>

4

5

6

1 **Appendix A: Tables**

2 **Table A.1.** Characteristics of the samples of monolingual and bilingual seniors tested in
 3 Experiments 1A and 1B. Mean values for each group are displayed with standard deviations
 4 between brackets. T-tests report independent sample t-tests comparisons' results.

5

	Monolinguals		Bilinguals		p value
Chronological age (years)	68.75	(4.62)	69.38	(4.59)	0.64
Education (age)	15.58	(3.15)	16.17	(3.86)	0.57
MMSE (raw score)	29.13	(0.99)	29.17	(1.17)	0.89
General IQ (percentile)	59.67	(31.27)	65.33	(29.25)	0.52
Spanish general	8.54	(1.02)	8.67	(1.17)	0.69
Spanish speaking	8.67	(1.05)	8.54	(1.06)	0.68
Spanish comprehension	8.75	(0.9)	8.79	(1.06)	0.88
Basque general	--	--	8.04	(0.95)	--
Basque speaking	--	--	8.13	(1.08)	--
Basque comprehension	--	--	8.29	(1.37)	--

6

7

8

1 **Table A.2.** Mean latencies for correct responses and error rates in all conditions for the tasks tested
 2 in Experiments 1A and 1B for the monolingual and bilingual groups of seniors. Reaction times are
 3 showed in milliseconds with standard deviations between brackets. Errors are showed in
 4 percentages with standard deviations between brackets.

5

		Verbal Stroop task (mean RTs)				Numerical Stroop task (Mean RTs)			
		Monolinguals		Bilinguals		Monolinguals		Bilinguals	
Conditions	Congruent	772	(217)	787	(167)	615	(133.2)	608	(111)
	Incongruent	1017	(202)	1001	(176)	693	(153)	684	(133)
	Neutral Word	901	(185)	892	(131)	621	(125)	631	(111)
	Neutral Symbol	791	(144)	780	(105)	--	--	--	--
	Total	871	(167)	865	(130)	643	(131)	641	(114)
Effects	Stroop	246	(167)	213	(128)	77	(86)	76	(60)
	Congruency	129	(164)	104	(128)	5	(37)	24	(38)
	Incongruity	-117	(104)	-109	(125)	-72	(90)	-52	(73)

6

		Verbal Stroop task (mean error rates)				Numerical Stroop task (mean error rates)			
		Monolinguals		Bilinguals		Monolinguals		Bilinguals	
Conditions	Congruent	0.52	(1.41)	0.87	(1.73)	0	(0)	0.26	(1.28)
	Incongruent	2.26	(4.07)	2.95	(5.83)	2.08	(5.43)	2.34	(4.81)
	Neutral Word	0.69	(2.01)	0.69	(2.01)	0	(0)	0.52	(1.76)
	Neutral Symbol	0.35	(1.18)	0.17	(0.85)	--	--	--	--
	Total	0.95	(1.53)	1.17	(2.09)	0.69	(1.81)	1.04	(1.94)
Effects	Stroop	1.74	(4.06)	2.08	(5.63)	2.08	(5.43)	2.08	(4.76)
	Congruency	0.17	(2.6)	-0.17	(2.29)	0	(0)	0.26	(5.04)
	Incongruity	-1.56	(2.96)	-2.26	(4.91)	-2.08	(5.43)	-1.82	(1.28)

7

8

- 1 **Table A.3.** Characteristics of the sample of bilingual seniors tested in Experiments 2A and 2B.
 2 Mean values are displayed with standard deviation between brackets.

	Bilinguals	
Chronological age (years)	69.36	(4.4)
Education (age)	17.71	(4.71)
MMSE (raw score)	29	(1.3)
General IQ (percentile)	70	(29.65)
Spanish general	8.72	(1.08)
Spanish speaking	8.65	(1.07)
Spanish comprehension	8.99	(0.99)
Basque general	6.49	(2.4)
Basque speaking	6.7	(2.62)
Basque comprehension	7.23	(2.13)

- 3
4

- 5 **Table A.4.** Mean latencies and error rates for the bilinguals seniors tested in Experiments 2A and
 6 2B. Reaction times are presented in milliseconds and error rates in percentages. Standard deviations
 7 are displayed between parentheses.

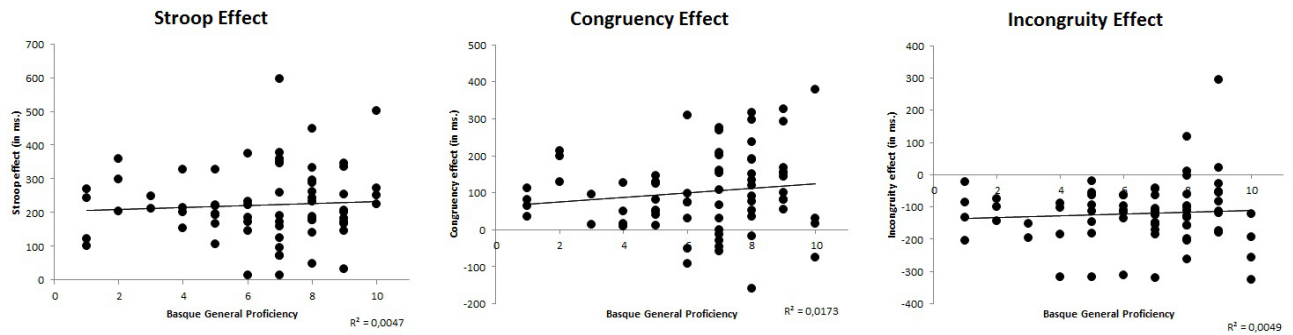
		Verbal Stroop task				Numerical Stroop task			
		Reaction Times		Error Rates		Reaction Times		Error Rates	
	Congruent	789	(145)	0.42	(1.26)	598	(103)	0.45	(1.62)
	Incongruent	1011	(174)	2.14	(4.42)	671	(117)	1.79	(4.00)
Conditions	Neutral Word	891	(140)	0.36	(1.37)	621	(100)	0.27	(1.28)
	Neutral Symbol	765	(104)	0.18	(0.85)	--	--	--	--
	Total	864	(128)	0.77	(1.46)	630	(103)	0.83	(1.60)
Effects	Stroop	223	(107)	1.726	(4.29)	73	(56)	1.340	(4.37)
	Incongruity	-120	(98)	-1.79	(4.11)	-50	(61)	-1.52	(3.90)

Congruency	103	(108)	-0.06	(1.66)	24	(34)	-0.18	(1.83)
------------	-----	-------	-------	--------	----	------	-------	--------

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11

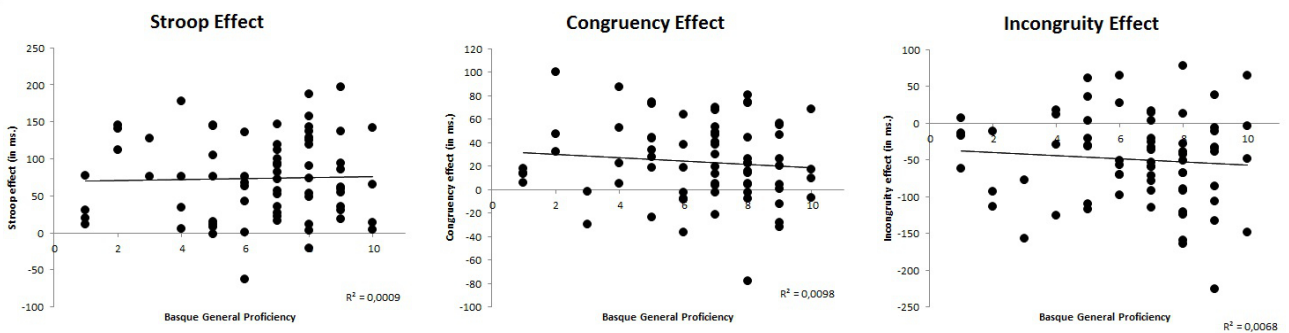
1 **Appendix B: Figures**

2 **Figure B.1: Correlations in the verbal Stroop task.** Correlations between *Stroop* (Incongruent-
3 Congruent), *congruency* (Neutral Words-Congruent) and *incongruity* (Neutral Words-Incongruent)
4 indices in latencies (in the vertical axis, indicated as the difference between conditions in
5 milliseconds) and General Basque Proficiency (in the horizontal axis, from 1 to 10) for the
6 Experiment 2A (verbal Stroop task). R^2 value is indicated in each graph.



7
8

9 **Figure B.2: Correlations in the numerical Stroop task.** Correlations between *Stroop*
10 (Incongruent-Congruent), *congruency* (Neutral-Congruent) and *incongruity* (Neutral-Incongruent)
11 indices in latencies (in the vertical axis, indicated as the difference between conditions in
12 milliseconds) and General Basque Proficiency (in the horizontal axis, from 1 to 10) for the
13 Experiment 2B (numerical Stroop task). R^2 value is indicated in each graph.



14
15
16